A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY

Military Operations Research

Fall 1995

V1, N3

A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY

MILITARY OPERATIONS RESEARCH SOCIETY

Military Operations Research

A publication of the Military Operations Research Society

The Military Operations Research Society is a professional society incorporated under the laws of Virginia. The Society conducts a classified symposium and several other meetings annually. It publishes proceedings, monographs, a quarterly bulletin, *PHALANX*, and a quarterly journal, *Military Operations Research*, for professional exchange and peer criticism among students, theoreticians, practitioners and users of military operations research. The Society does not make or advocate official policy nor does it attempt to influence the formation of policy. Matters discussed or statements made in the course of MORS symposia or printed in its publications represent the opinions of the authors and not the Society.

Note from the Publisher:
Beginning with this issue, members of MORS will receive the journal with no further charge. Service sponsors will cover the cost for some staff members. All others must purchase a subscription at the rate of $40.00 for one year or $75.00 for two years.

We hope you find this issue of interest and use to you. As always, the Board of Directors and Staff of the Military Operations Research Society ask for your feedback on this or any of our other publications or programs.

—*Natalie S. Addison, Publisher*

 ISSN Number 1082-5983

Printed on Recycled Paper

Table of Contents

Volume 1, Number 3

Fall 1995

64th MORSS Working & Composite Groups

Theme: *Leveraging Technology for the Military Analyst**

Framing the Analysis ✦ Organizing the Data ✦ Exercising the Tools ✦ Conducting the Analysis

Working/Composite Group Coordinator:	**Kerry E. Kelley, USTRATCOM/J502**	**402-294-4102**
CG I: STRATEGIC	**Pat McKenna, USSTRATCOM/J533**	**402-294-5308**
WG 1: Strategic Operations	Dr. Gene Schroeder, LANL	505-665-3101
WG 3: Arms Control & Proliferation	Dr. Robert Batcher, ACDA	202-736-7396
WG 4: Revolution in Military Affairs	Thomas Mahnken, OSD Net Assessments	617-576-0124
CG II: NAVAL WARFARE	**Bill Mulholland, McDonnell Aircraft Company**	**314-232-9647**
WG 5: Expeditionary Warfare/Power Projection Ashore	Timothy J. Sullivan, Texas Instruments	214-575-5642
WG 6: Littoral Warfare/Regional Sea Control	CDR Kirk Michealson, N815L1	703-697-8078
CG III: AIRLAND CONTINGENCY OPERATIONS	**James L. Wilmeth III, SETA Corporation**	**703-695-4657**
WG 2: Missile Defense	Beverly Nichols, USA SDC	703-602-7124
WG 7: Nuclear Chemical Biological Defense	MAJ Jerry Glasow, HQDA-DCSOPS	703-697-5752
WG 8: Mobility	Denis Clements, GRC International	703-607-3192
WG 9: Air Warfare	Thomas Lillis, McDonnell Aircraft Company	314-234-2737
WG 10: Land Warfare	Larry Cantwell, TRAC	913-684-6867
WG 11: Special Operations and Operations other than War	August (Greg) Jannarone, Consultant	813-677-9498
WG 12: Air Defense	Michael Minnick, Lockheed Martin	609-722-7741
WG 13: Electronic Warfare and Countermeasures	LtCol Tom White, USAF	210-977-2427
WG 14: Joint Campaign Analysis	Richard P. Morris, McDonnell Douglas Aerospace	703-412-3944
CG IV: SPACE/C3I	**John R. Pettit, NSA**	
WG 15: Command Control and Communications	William Kemple, NPS	408-656-2191
WG 16: Military Environmental Factors	Eleanor Schroeder, Naval Oceanographic Office	601-688-5502
WG 17: Operational Contribution of Space Systems	TBD	
WG 18: Operations Research and Intelligence	Peter A. Shugart, USA TRAC	505-678-2937
CG V: RESEARCH AND DEVELOPMENT	**John M. Green, Lockheed Martin**	**609-722-4516**
WG 19: Measures of Effectiveness	Robert Meyer, NAWC	619-927-1279
WG 20: Test and Evaluation	Michelle Kirstein, HQ AFOTEC	505-846-1938
WG 21: Unmanned Systems	CPT Edward Kleinschmidt, USMA	914-938-5664
WG 22: Cost and Operational Effectiveness Analysis	LTC Robert Clemence, Jr., OSD(PA&E)	703-697-1600
WG 23: Weapon System Acquisition/Requirements Analysis	Terrence Cooney, Veda, Inc	513-476-3506
CG VI: RESOURCES AND READINESS	**Mary JoAnn Carroll, AFSAA**	**703-697-1622**
WG 24: Soft Factors in Military Modeling and Analysis	William M. Pugh, Naval Health Research Center	619-553-8403
WG 25: Social Science Methods	Dr. Jock Grynovicki, ARL	410-278-5917
WG 26: Logistics	Alan Cunningham, USA TRAC	804-765-1830
WG 27: Manpower and Personnel	Dr. David M. Rodney, CNA	703-845-2346
WG 28: Resource Analysis and Forecasting	Dr. Thomas Frazier, IDA	703-845-2132
WG 29: Readiness	John E. Leather, DMDC/TREAD	408-583-2400
CG VII: METHODOLOGIES AND TECHNOLOGIES	**LTC Michael McGinnis, USMA**	**914-938-5528**
WG 30: Decision Analysis	Samuel (M.att) Vance, McDonnell Aircraft Company	314-232-9747
WG 31: Computing Advances in MOR	CPT Neal Lovell, TRADOC Analysis Ctr	408-656-3086
WG 32: Advanced Analysis, Technologies and Applications	Mark Axtell, VEDA, Inc.	513-253-4770
WG 33: Modeling Simulation and Wargaming	Steve Packard, Oak Ridge National Lab	615-574-9388

*See Request for Application on Page 26.

"Graphing" an Optimal Grand Strategy

John Arquilla
Associate Professor of National Security Affairs

Hal Fredricksen
Professor of Mathematics

Naval Postgraduate School
Monterey, CA 93943

ABSTRACT

Graph theory provides a useful framework for generating insights into problems of sufficiency and optimality across a wide variety of physical relationships. Applied to the realm of grand strategy, this approach assists in developing a methodology for estimating the minimum level of forces required and determining the optimal deployments for the successful pursuit of national security goals. In theory, the adoption of a defense-in-depth maneuver strategy provides the most efficient use of scarce resources. However, deterrence stability attenuates due to the absence of robust local balances of forces. Comparative case analyses of the Roman and British empires confirm the efficiency of depth defense, as well as the weakening of deterrence. Implications for U.S. policy are that, despite sizeable reductions, two regional wars can be fought and won, nearly simultaneously, even below base force levels. However, the deployments required to effect this grand strategy may make challenges to conventional deterrence more likely. Finally, it is demonstrated that small increases in forces above minimum requirements create a valuable "margin of safety" and may significantly improve crisis and deterrence stability.

The specter of decline confronts all great powers eventually. A substantial body of literature, associated generally with theories of either power transition or cycles of relative power, addresses the onset of and efforts to cope with this unavoidable problem (Doran 1971; Organski and Kugler 1980; Gilpin 1981; Kennedy 1987; Modelski 1987; Goldstein 1988). Fundamentally, decline poses a strategic dilemma: that of either trying to maintain the status quo with scarce resources, even by means of preventive war (Gilpin 1981, 191); or by retrenching, unilaterally reducing spheres of influence.[1] Both approaches can entail great risks.

Imperial Spain, for example, faced with declining resources, attempted to hold all of the vast gains it made in the 16th century, and found itself consistently "overstretched," unable to deter predatory attacks, or to defend successfully against them (Elliott 1991). The Soviet Union, on the other hand, recognized its material deficiencies, and chose, a few years ago, to retrench preemptively, resulting not only in the swift breakup of its imperium, but also in the substantial dissolution of its own polity.

Of the two strategic approaches, "holding the line" appears less risky at the margin, perhaps because the effects of decline may be mitigated by spreading them over a longer period of time. The Spanish empire took nearly 300 years to collapse, from the loss of Holland in 1609 to the war with the United States in 1898. The Ottoman empire followed a similar temporal pattern of senescence. The former Soviet Union, which instead chose retreat, is absorbing the substantial, wrenching consequences of imperial loss over an extremely short period. Even Britain, which withdrew skillfully from empire in the wake of the Second World War, suffered some of the immediate economic and politico-military consequences of strategic retreat, though they were cushioned by the willingness of the United States to fill the British void.

Because of the seemingly high risks of retreat, even of partial withdrawal, this study concentrates its analysis on the stodgier option, maintaining the status quo, and upon the implications, for deterrence and defense, of adopting a grand strategy of "holding the line." It introduces first a mathematical graphing methodology by means of which optimal choices for using scarce resources may be identified. Previous efforts along these lines have employed insights from geometry to develop a framework for successful defense at the tactical and operational levels (Gupta 1993). This study focuses upon the grand strategic level of analysis, and considers the prevention of war as well as defense against aggression. With regard to successful deterrence, special attention is given to the perceived need to achieve favorable immediate or short-term local balances of forces (Huth and Russett 1984, 1988; Huth 1988).

This new methodological approach to optimizing security strategies is then applied retrospectively to the two most notable historical cases of great empires, one primarily continental, the other maritime, as they confronted potential decline: Rome in the 4th century and Britain before World War I. Insights drawn from these studies will then be applied to help analyze the current situation of the United States in the post-Cold War world, with special emphasis given to the propriety of the current

American strategy of attempting to fight and win two regional wars nearly simultaneously.

OF GRAPHS AND "PEBBLES"

More than any other formal analytic framework, graphical analysis has long been effectively used to model a wide variety of physical relationships (Kemeny and Snell 1962), from the feeding habits of rainforest animals (Harrison 1962) and the optimization of municipal services (Tucker 1973), to the most efficient means for assigning radio frequencies (Hale 1980).

This study views regions of the globe as nodes on a graph. An edge between nodes indicates their proximate relationship. Two nodes are adjacent if they are incident on the same edge; and any two nodes that are not adjacent are called separated. The distance between adjacent nodes is exactly 1. The distance between separated nodes is equal to the number of edges between adjacent nodes on the shortest path that connects them.

Further, "pebbles," notional representations of the minimum forces needed in order to control a node, are introduced and placed on the nodes of the graph. The pebbles may be moved from node to node along edges. Such moves on the graph are reminiscent of, and have their basis in, classical games such as Awari, an offshoot of the Mankale'h family of games, including backgammon. In these ancient games, pebbles on a node may be swept up and distributed among adjacent nodes by means of movement along incident edges.

In this game, a pebble can move from one node to another; however, a cost of "1" is assessed for each pebble moved. Additionally, no pebble may move if it is the last pebble remaining at the node. With regard to lengthier moves, first k-1 pebbles are moved to an adjacent node along the path to the terminal node. The cost of this initial move is k-1, as there were k-1 pebbles moved along the first edge. Then k-2 pebbles are moved along the second edge, k-1 along the third edge, etc.; until, finally, the last pebble is moved along the k-1st edge to the terminal node. The resulting structure is a string of k adjacent nodes, each of which possesses exactly 1 pebble and is connected by edges. This forms a "bridge" of length k. The cost to build the bridge is the total cost of all of these moves. Cost is given as:

Eq. 1: (*Cost Determination*)

$$1+2+3+\ldots+k-1 = k(k-1)/2$$

The goal in moving pebbles is to place one at any prescribed node from a given initial configuration. If this can be accomplished, the original deployment of pebbles achieves the sufficient condition to qualify as a winning situation. Each pebble, therefore, represents a force able to seize, hold and control the territory represented by the nodes. Thus, a single pebble comprises all that is necessary, to a high degree of probability, to win a regional conflict at that location.[2] It is assumed, for analytic purposes, that the composition of pebbles does not include allied forces. Contingency and military requirements planning must consider the possibility that a nation will have to fight alone on occasion, or that allies might prove either inconstant or ineffective.

The bridge that is built to produce a win is consistent with notions of the strategic importance, and costs, of secure lines of communication, and has analogs in history. The closest parallel, no doubt, may be seen in the island-hopping strategy employed by the United States in the Pacific War against Japan (1941-1945), wherein "bridges" were built from one staging point to the next, with each providing security for the advance of further forces. Even in the recent Gulf War, where the Mediterranean was simply used for transit, significant forces had to be employed to keep the "bridge" secure at all times (Owens 1995, 80).

The more distant the target node, the greater the cost of reaching it in terms of time and resources expended. In this model, no further costs, such as placing a cost label on edges between nodes, or intra-nodal movement, are assessed. Even so, the costs of bridge-building are substantial. However the existence of the bridge has great importance in terms of developing the capability for reconcentration of forces to meet a second conflict that arises even while the first is in progress, a scenario envisioned by the current U.S. warfighting doctrine of retaining the ability to fight two regional wars with near simultaneity. The costs described herein reflect

reasonably the concept of the loss-of- strength gradient (LSG) described by Boulding (1962).[3] Simply put, this formulation holds that the projection of power entails predictable costs that increase, in the aggregate, with distance.

The initial placement of pebbles can be crucial to producing a win without incurring unreasonable costs due to having a large number of required moves. Further, this study assumes that deterrence of aggression at any specific node can be achieved by an initial placement of pebbles that allows a win at less than a cost of 2; that is, if there is a pebble on the node in question, or if one may be deployed there at a cost of 1. This captures the theoretical notion that deterrence success depends heavily upon the immediate and short-term local balance of forces in a crisis (Huth and Russett 1984, 1988; Huth 1988). It also recognizes that, while bridge-building may provide, ultimately, an adequate form of defense, it nevertheless may not deter well in those cases where response times are slowed by distance factors.

These distance factors imply a need to distinguish between "simple" and "complex" graphs. The former possess a node adjacent to all others, the latter do not. Further, there are levels of complexity, measured in terms of the number of nodes at a distance of "2" from any given node. The key point introduced by this concept is that complex graphs contain the eventual, but inevitable need for bridge building. Thus, complexity entails varying costs for alternate initial deployment schemes, which can differ radically. In particular, when the need arises for movement to control a second key node, graph complexity greatly complicates optimality calculations. One means for dealing with complexity when contemplating initial deployments, we hypothesize, is to break down complex graphs into simple components, assuming that this approach will encourage deployments designed to minimize the costs of graph coverage.

Current U.S. defense strategy, for example, calls for placement of sufficient force to allow for winning two regional wars nearly simultaneously (Aspin 1993). In this instance, sufficient pebbles must be initially placed so that a win can be achieved at any given location, followed by a win at any second location. Note that pebbles moved to produce the first win at the location of the initial conflict cannot typically be moved, as they will now be single pebbles incapable of movement on their own. However, these pebbles do provide a bridge along which any individual pebble may travel. In this case, the cost of using the established bridge is just 1 for each step along the way. Then, the cost of the original configuration is computed as the cost to achieve the win-win scenario. This is the worst-case cost for the choice of any ordered pair of nodes involved in the win-win scenario, where the costs of

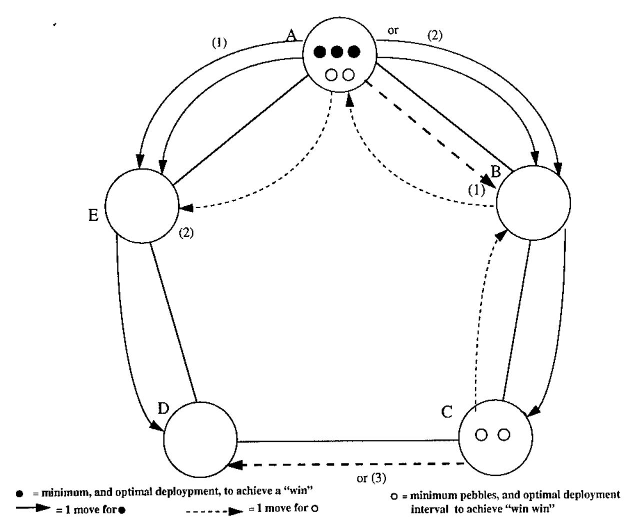


Figure 1. The "Pentagon"

producing the respective wins at the two nodes is minimized by following the shortest possible paths from nodes pebbled in the original configuration.

A THEORETICAL EXAMPLE

To illustrate the concepts and definitions developed above, one may consider the cycle graph on five nodes (C5, see Figure 1). No particular meaning need be associated with the nodes of this pentagonal structure, though they could easily be related to areas of interest to the United States: the Western hemisphere, Europe, the Near East, South Asia and the Far East. The letters A through E are assigned to the five nodes, beginning at the top of the pentagon and proceeding in clockwise order. With regard to the issue of "complexity," the pentagon's is equal to 2, as there are two nodes at distance 2 in the graph for any given node.

From the rules of movement, it should be clear that possession of either one or two pebbles will not allow for achieving even a single "win." Therefore, our illustrative analysis begins with three pebbles. To place three requires, initially, the use of 1, 2, or 3 locations. Though every possible configuration is analyzed herein, one must also remain cognizant of the practical limitations that might develop in reality, as economic constraints (unwillingness to pay host-costs) or political sensivities (such as to having non-Muslim troops permanently stationed on the territory of an Islamic state) may prevent deployment to certain nodes.[4] There is only one way that three pebbles can be assigned to a single node, and two ways that two nodes may be used. These represent the two-fold choice of whether to select nodes that are adjacent or separated. Distractions of the left or right are not considered as being distinct of the two nodes where one gets 2 pebbles and one receives just 1. If three nodes are utilized, there are again just two possibilities, either all contiguous, or one separated from the others.

The latter two cases can be ruled out quickly, and may be illustrated by the 5-tuple configurations (11100) and (11010) respectively. In both cases three nodes are covered with no moves permitted, failing to achieve a win. When two nodes are utilized initially, a single pebble move may be made from the node containing 2 pebbles. This may be depicted as (20100)—>(11100)-1, or (20100)—>(10101)-1. The arrow indicates a move, and the nonparenthetic integer reflects accumulated costs. In total, four different nodes have been covered at a worst case cost of 1. No winning configuration appears. The case of adjacent nodes is illustrated as (21000)—>(11001)-1, or (21000)—>(12000)-1—>(11100)-2. The cost is one for the first move in either case, but a second move is possible at an additional cost of 1. Again, exactly four nodes are covered, now at a worst case cost of 2. Still, no winning configuration emerges from these placements.

Where all 3 pebbles are placed at a single node initially, there are more possibilities. Now, two pebbles can move to an adjacent node, and then a single pebble can be moved further. This is illustrated by (30000)—>(21000)-1 —>(12000)-2—>(11100)-3. Here, the worst-case cost is 3 to cover three nodes. Since the moves may also be made in the counterclockwise direction, the situation may be alternately rendered as (30000)—>(20001)-1—>(10002)-2—> (10011)-3, again covering three nodes at a cost of three. Together, at a worst-case cost of 3, all five of the nodes (A-E) have been covered in one sequence or the other; and a winning configuration is established by virtue of all the nodes being covered via this "depth defense" deployment of the pebbles. (30000) is a winning configuration whose cost = 3.

Though a winning configuration emerges at the 3-pebble level, no "win win" scheme has yet been identified. Therefore, the analysis must proceed to the level of 4 pebbles. In this scenario, only the case of the initial occupation of four nodes (11110) fails to achieve a winning configuration. In all other cases, a win with 4 pebbles is easily established. Costs for each "single win" option are given in Table 1.

There are only two ways, though, to establish a "win win" scenario with an initial placement of 4 pebbles. If all are placed at a single node, then regardless the order in which two nodes are named, it is possible to reach both. Either adjacent node is reached in 1 step: (40000)—>(31000)-1, or —> (30001)-1. Then, either (originally) separated node can be reached utilizing one or the other bridge in two additional steps, for a total cost of 3. If one of these separated nodes were the first named, then it could

Table 1. Winning 4-Pebble Configurations.

COST	CONFIGURATION
1	(20110), (20200)
2	(21010), (21001)
3	(21100), (30100), (31000), (40000)
4	(22000)

be covered in 3 steps: (40000)—>(31000)-1—>(22000)-2—>(21100)-3. Then, if an originally separated node is named second, it can be reached in three additional steps utilizing an established bridge: (21100)-3—>(12100)-4—>(11200)-5—>(11110)-6. The win-win scenario is thus achieved for this configuration (all pebbles on one node), and the cost for the double win = 6.

The other way to achieve a win-win is by means of the placement (20200). The first chosen node (or conflict) can be reached at a cost of 1 in every case. A second occurrence anywhere may then also be reached, though the cost varies with the location. Cost analysis requires the consideration of two cases. The first named node may be adjacent to just one of the initially occupied nodes, in which case the second node can be reached in just one more step, for a total cost of 2. Alternately, the first named node could be adjacent to both occupied ones. In this instance, referring to Figure 1, one might hypothesize 2 pebbles on A and 2 on C, with a conflict erupting on B. Either choice of how to occupy that node (moving from A or from C) establishes a bridge (20200)—>(21100)-1, to the most distant node, which can now be reached in three additional moves (21100)-1—>(12100)-2—>(11200)-3—>(11110)-4. The worst-case cost to achieve the win-win scenario = 4.

With a minimum number of pebbles needed to achieve a double win of 4, it is interesting to engage in some marginal analysis, to consider the effect, in terms of cost savings, of adding one pebble, for a total of 5. In this case, though, placement on only one or two nodes is considered, reflecting implicitly the political or economic limitations on "overseas" basing that often arises in reality. Given this restriction, only five cases require consideration: (50000), (41000), (40100), (32000) and (30200). Win-win scenarios exist in every case, but the costs are startlingly different. In the respective cases described above, the costs are 6, 5, 4, 4, and 2. Thus, in light of the best deployment of 5 pebbles at two or fewer nodes, *a modest increase in force size may generate cost savings of 50%*. That is, cost computed in terms of the required time to respond effectively to any threat. For deterrence stability, such a finding could have profound implications.

The best-case scenario plays out as (30200)—>(21200)-1, and then the second named node can be covered by one step in either direction, without requiring the use of the bridge. Of course, the costs of increasing force size by one-fourth might generate domestic political resistance, if the only cost saving were a one-half reduction in response time in crisis. On the other hand, if such a marginal increase cut the likelihood of an outbreak of war, or other failure of deterrence, in half, then the gains might well be viewed as greater than the costs.

The foregoing theoretical examples form the basis for the following case studies, which key on force minimization and optimization, for deterrence and defense. The situations faced by Rome and Britain, when their empires were territorially most expansive and their material constraints were beginning to be sharply felt, have been chosen for comparative analysis. It will also be argued that the United States faces a similar situation in the post-Cold War world, with still-extensive commitments and interests coming into tension with economically straitened circumstances.

HISTORICAL CASES

The Pax Romana

At the height of its power in the 3rd century, Rome fielded approximately 50 standing legions, aggregating over 300,000 combat troops, exclusive of auxiliaries (Gibbon [1776]1937; Luttwak 1976; Macmullen 1980; Delbruck [1921]1990). These forces deployed evenly to the many border areas of the imperium, providing Rome with

an explicit "forward defense" based on a doctrine of positional warfare. Additionally, their presence at the marches of the empire encouraged friendly barbarians to settle nearby, creating an "external depth" that further enhanced security (Ferrill 1991). Over time, though, the empire stagnated economically, and couldn't maintain the same large military machine that characterized its halcyon days.

Thus, by the 4th century, the Roman armed forces had shrunk nearly in half, from 50 to roughly 25 standing legions (Luttwak 1976, 189). Forward, positional defense grew infeasible. At this point, Constantine instituted quantum changes in Roman grand strategy, going well beyond the "shallow depth" hybrid defensive scheme advanced by Diocletian over a century earlier. Constantine directed the redeployment of the legions, roughly half in the environs of Rome and Northern Italy, the remainder in the vicinity of the city which bears his name. The defense-in-depth maneuver strategy that underlay this shift held that the legions would now be better suited to responding to any emergent threats around the perimeter of the empire (Jones 1964; Luttwak 1976). This approach conceded that some territory would be lost in the initial phase of conflict, but that it would be reconquered expeditiously.

For purposes of graphical analysis, each Roman "pebble" consists of six legions of regular field forces, or *comitatenses*. This figure derives from Gibbon's and Delbruck's accounts of the forces needed for the major campaigns of the 2d-4th century period. Of course, exact numbers of troops remain shrouded in obscurity, but the numbers of legions engaged often emerges with clarity. Again, as in the previous theoretical example, allied forces and barbarian levies, because of their unreliability, so well borne out historically, are not included in the "pebbling" process.

Figure 2 depicts the strategic schematic of the Roman empire at its height, with key nodes set at the two capitals and all perimeter areas. Edges represent, and are in line with, Rome's principal roads and maritime lines of communication. The compexity of this graph is three. Using the six-legion pebbling guideline generates eight pebbles in the 2d century, four in the 4th. Thus, at its greatest strength, the empire could continuously maintain enough forces to win a field campaign on each of its key nodes. The legions were indeed "legion." Two centuries later, though, four pebbles had to cover the same eight nodes. Was the Constantinian depth defense the correct answer for Rome, as Luttwak (1976) has argued? Was it optimal?

Constantinian depth defense may be represented, using Figure 2, by envisioning the placement of two pebbles each on Rome and Constantinople. This configuration does allow for the achievement of a "win." That is, these

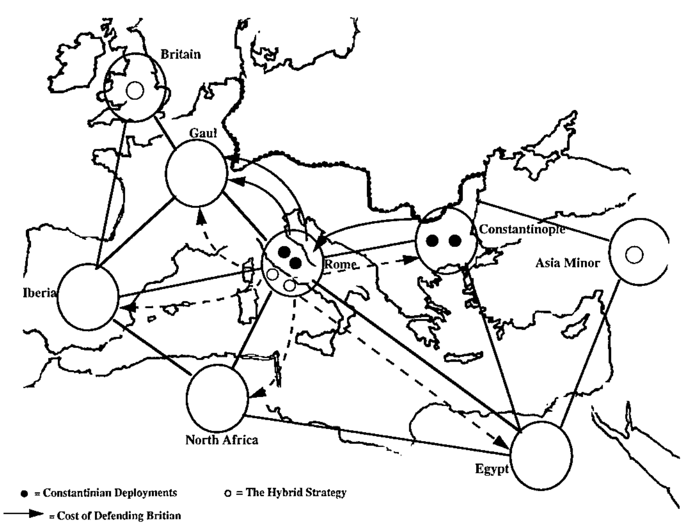


Figure 2. Strategic Schematic of The Roman Empire

forces may deploy to any region in imminent crisis or war. Response time to all locations, save for Britain, is 1 each, implying that deterrence ought to hold in these areas. Reaching Britain, however, entails a cost of 4, as a Rome-Gaul bridge must be established (1) and then a pebble can move from Constantinople-Rome-Gaul-Britain (+3, total cost of 4). If an edge existed between Gaul and Constantinople, the cost of moving to Britain would reduce to 3, but the interposition of the unfriendly, and unconquered Germans made this infeasible.

Thus, the total cost of Constantine's depth defense is 9, and the strategy does not guarantee a "win win" situation either, although Britain is the only node that prevents this. The difficulties and costs of defending Britain had long concerned Roman leadership, leading early on to the creation of Hadrian's Wall, the Maginot Line of its day. Presumably, this fortification acted as a deterrent and force multiplier, freeing up the legionary field forces for the defense of the rest of the imperium. Interestingly, there are a variety of lower cost solutions to the Roman strategic dilemma.

The first alternative strategy would keep 2 pebbles in Rome, but only 1 in Constantinople. The 4th would deploy initially to Gaul. This covers one more node initially, improving deterrence, and reduces the cost of movement to Britain to 2. On the other hand, responding to a crisis on the frontier with the Persian Empire now costs 2. Total costs are 7 under this strategy. When win-win considerations are added, though, this strategy would leave 4 nodes beyond reach in the event of a second war, as opposed to the 1 (Britain) that would be sacrificed in the actual Constantinian strategy.

Obviously, political importance was attached to the forces deployed in and around Rome, and it would be hard to conceive of a strategy that failed to retain substantial forces there. Nevertheless, for analytical purposes, it is interesting to consider a second strategic approach, with two variants, in which Rome alone, then both Rome and Constantinople are neglected. In the first variant, 2 pebbles each are placed in Gaul and Constantinople. The cost is 1 to move to all nodes, except for North Africa, which has a cost of 3. Thus, the aggregate cost is 8, marginally better than the actual strategy employed. With regard to fighting two conflicts, this strategy leaves two nodes vulnerable, one more than under Constantine's strategy; unless the first conflict occurs in North Africa, after which response to second conflicts grows problematic. Thus, this strategy would likely require a willingness to sacrifice North Africa.

For example, a first war in Britain leaves Iberia and North Africa open after the initial Gaul-Britain move. Similarly, a war in Asia Minor leaves Egypt and North Africa beyond reach. Finally, a war in Egypt would result in uncovering Asia Minor and North Africa. The problem with the Gaul deployment is that it makes bridge building to reach the second conflict impossible (except if the first war is in Rome) when a bridge is required to move field forces to distant regions.

The most effective "non-Rome" option for coping with the abovementioned situation would be to keep 2 pebbles in Gaul, but to shift the 2 in Constantinople to Egypt. This reduces total costs to 6, a one-third savings over the historical strategy employed. However, the problem of covering the second conflict is double that of the Constantinian strategy. For example, a first war in Asia Minor leaves Constantinople and North Africa uncovered. If the first conflict is in the west, though, say Iberia or Britain, then only one node remains uncovered. Thus, this strategy would have reduced the overall costs of winning substantially, and only marginally worsened the empire's ability to respond to a second conflict. Of course, the dark political consequences of removing forces from the two capitals are hard to contemplate and would, undoubtedly, have been politically unacceptable.

With this last concern in mind, returning pebbles to Rome, aside from its political attractiveness, allows for introduction of the most efficient "one war" strategy. This consists of a hybrid of forward and depth defenses, with 1 pebble each in Britain and Asia Minor, and 2 in Rome. In this configuration, the overall cost of dealing with one war anywhere is 5. Unfortunately, the response to the first war leaves four nodes open that cannot be protected in the event of a second conflict. Adding 1 more pebble in Rome would solve this problem, but such an increase would have strained the empire greatly.

On balance, the Constantinian depth defense strategy forms the most attractive option. Its seemingly high cost in winning one war is an artifact of the cumbersome moves required to defend Britain, where this analysis implies deterrence would most likely fail. On the other hand, for purposes of coping with two wars, this strategy minimizes the vulnerable nodes (to one, Britain), a better result than any other deployment scheme. No historical record suggests that the Romans employed graphical analysis in their strategic planning; but it is interesting that they identified Britain as a particular defensive problem, leading them to the creation there of one of history's more ambitious efforts in fortification.

If the Romans demonstrated such refined strategic insight, then what accounts for the collapse of the empire? Ferrill (1986) argues convincingly that the fall of the empire resulted from a lessening of the military effectiveness of the legions. The move to depth defense caused a reshaping of the legionary forces to incorporate more mobile cavalry which, Ferrill argues, inadvertently debilitated the Roman infantry by sapping its numbers. Gibbon ([1776]1909, 322) contends that the fall of the empire came soon after Roman infantry discarded their body armor (so that they could move faster on the march).

Thus, Rome fell because its "pebbles" lost their value of "1." They increased mobility at the cost of fatally compromising their hitting and holding power. This problem does not appear to apply to the next, British, case; but it resonates with current debates about the appropriate "lightness or heaviness" that U.S. forces should possess in the post-Cold War world. In another respect, the Roman case also compares more closely with the United States than Britain, in that Rome combined land and sea power, while Britain remained essentially a maritime power, albeit one that could, when necessary strike powerfully on land. Finally, the British case will prove somewhat clouded, because of the presence of a rapidly emerging rival (exemplified by the German maritime threat). Rome in the 4th century faced a multitude of smaller, potential threats, much as the United States does in the immediate post-Cold War period.

The Pax Britannica

Like Rome, Britain enjoyed a period in which it was a sole "superpower." In 1816, a year after the fall of Napoleon, the Royal Navy possessed two-thirds of the capital ships *in the world*. It outnumbered the Russian fleet, in terms of warships of 60 guns or more, by 5:1; France by 4:1; and the United States by over 20:1. Additionally, Britain enjoyed almost equally favorable margins in smaller vessels such as frigates and gunboats (Modelski 1988). For comparative purposes, though, in terms of setting the number of "pebbles," only British capital ships are considered, as were only regular Roman legions in the previous case. Smaller vessels are analogous to the Roman auxiliaries, and equally vulnerable. In the words of First Sea Lord Jacky Fisher, any moderate-sized opponent could "lap them [all] up like an armadillo set loose on an ant-hill!" (Kennedy 1976, 216).

In addition to its overwhelming naval mastery in the immediate post-Napoleonic period, Britain also enjoyed control of three-fourths of world trade. However, the Pax Britannica itself, which ensured the freedom of the seas and fostered trade openness, spurred growth at such rates that Britain itself could not avert a tendency toward relative decline. By 1860, its share of world trade had fallen to 25%, by 1900 it was 17% (Kennedy 1976, 190). In terms of naval power, its share of global capital ships fell from nearly 70% in 1816 to 50% in 1860 and, in 1900, to under 40%.

Like the Roman legions, the capital ships of the Royal Navy, early on, deployed to the various perimeters of the empire, providing a robust forward defense. For example, in 1848, four-fifths of the warships of the Royal Navy were stationed outside the home islands (Kennedy 1976, 170). In terms of "pebbling," where six Roman legions comprised a force capable of dominating any given sector, the British Admiralty appeared to have followed a loose policy of amassing twenty-ship battle fleets during the age of sail, reducing this to eight in the era of steam propulsion and steel. These minimums needed for local dominance would give Britain an inventory of six "pebbles" in 1816, reducing to four by the turn of the 20th century. In this British case, the small professional army

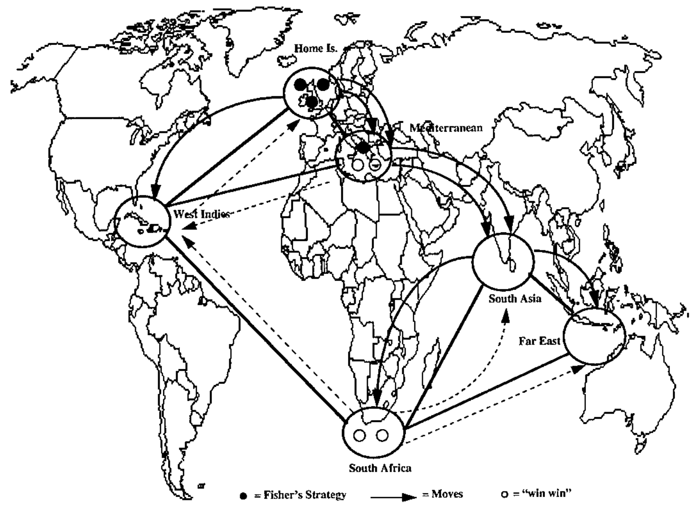


Figure 3. Strategic Schematic of the British Empire

is held constant, and is considered as a "projectile" fired by the Royal Navy into any given regional conflict; one, when properly supplied and supported, which consistently provided the empire with a high probability of success.

Figure 3 presents the British strategic schematic, with nodes at all key reaches of the empire, which has a complexity of 2. Early in the 19th century, the Royal Navy had the wherewithal to remain on station with dominant forces at each of the six key nodes of the empire. By 1900, though, down to four pebbles, small adjustments had been made. Instead of one pebble per node, two were kept in home waters, with one each in the Mediterranean and one in South Asian waters. This still had much of the appearance of a perimeter positional defense, or at least was no better than Diocletian's "shallow depth" strategy for the Roman legions. In 1905, though, Admiral Fisher, driven by his fear of dispersing his scarce resources in the face of rising German naval power, moved to a full depth defense. In schematic terms, he placed three pebbles in home waters, leaving one in the Mediterranean, and none elsewhere (Fisher 1919; Marder 1940; Kennedy 1976, 217).

Was Fisher right? Was his redeployment efficient? Optimal? The Fisher strategy of defense-in-depth could see to it that British forces arrived anywhere eventually. To achieve winning conditions against any one adversary, the "pebbling cost" of his approach amounts to 11, with 5 of this cost accounted for by movement to the Far East, 3 to reach the Cape of Good Hope and 2 to respond in South Asia. With regard to coping with a second conflict, Britain could reach any second war as long as the first one developed in South Asia. This would create a bridge which could then be employed to reach either the Cape or the Far East.

The high costs of the Fisher strategy, and the conditionality of the "win-win" capability encourage some application of graphical analysis in pursuit of more efficient strategies. The central problem to tackle in the British case is the great distance between Britain and the Far East. One solution consists of reducing the pebbles in the home islands to 2, and placing 2 in the Far East. This allows for a response time of 1 to any conflict, with an aggregate cost of 4. The cost is the same if the Far East forces are stationed in either South Asia or at the Cape of Good Hope. Alternately, 2 pebbles each in the Mediterranean and at the Cape, or in the West Indies and South Asia, also have aggregate costs of 4. These latter two cases have the additional benefit of creating win-win situations, as they permit a successful response to any second conflict that arises, at worst-case additional costs of 3.

Given the strategic imperative of balancing against the rising German threat, Britain would likely not have adopted either of the win-win

strategies, as each requires weakening the defenses of the home islands. Given the incredible popularity of the "invasion scare" literature of this period, there can be no doubt of the domestic political constraints upon any effort to move the fleet to peripheral areas. However, it might have been possible to reduce the home fleet to 2 pebbles (from Fisher's 3). Then, with 1 in the Mediterranean and 1 in South Asia, a winning situation could be achieved at a total cost of 7, a 44% saving on the Fisher plan. Nevertheless, this scheme fails to achieve the conditions for a double win.

To summarize, graphical analysis points out that Britain could have achieved a double win, and reduced its single win costs by 80%, by moving its fleet units to peripheral areas. This finding suggests that, under some circumstances, defending forward, even when limited by straitened economic conditions, may prove superior to a pure depth defense. Aside from cost reductions, of course, this approach maintains the kind of forward presence needed to shore up deterrence. As the American case unfolds, this insight about deploying forward may have particular value, as the United States suffers nothing of the sort of great power rivalry, or vulnerability to invasion, that Britain confronted, or thought it did, a century ago.

Implications for the United States in the Post-Cold War World

Can this graphing methodology, along with insights from the foregoing historical cases, help to determine the appropriate size of and optimal deployment scheme for U.S. forces? Or, more simply put, how many pebbles are needed, and where should they be placed? These questions may be answered explicitly, though the issue of quantitative requirements depends upon the definition of the force size necessary to achieve a probable win in any regional conflict. How much is an American pebble? For Rome, it was six legions of regular forces. For Britain in the age of steam and steel, it was eight capital ships, their supports, and an expeditionary force. An American pebble, however, must be more multifaceted, comprising sea, air and ground forces.

Perhaps the best guidance to determining the specifics of an American force sufficient to win a regional conflict comes from the "bottom-up" study prepared by Les Aspin. In it, the former defense secretary argues first that, in order to win one regional conflict, the United States would have to maintain, at the upper level, standing forces of 5 divisions of ground forces, 10 air wings and 5 carrier battle groups (1993, 10). These figures include what is needed for standing treaty commitments (in Europe and Northeast Asia, principally) and for purposes of the rotation base (mostly in the continental United States). His analysis does not specify the amounts needed for the "win win" strategy, but implies that doubling the force is not quite necessary.

Accepting Aspin's approach, one may then conservatively estimate total force requirements of 10 divisions, 20 air wings and 10 carrier battle groups, an amount that comes very close to the revised "base force" projections with which his study concludes (1993, 17). In actuality, though, U.S. forces will likely have a few less air wings and at least two more carriers. Since each carrier services approximately one wing of attack aircraft, the net figures remain unchanged.

Extrapolating from Aspin's figures, one may now configure each U.S. pebble as composed of 3 divisions of ground forces, 3 carrier battle groups (including a Marine expeditionary force) and 5 air wings (approximately 375 aircraft). Reserves are not included in these figures, as their impact is felt most substantially in longer wars, given their need for mobilization and combat training.[5] Regional wars will remain the province of regular American forces, much as regular Roman legions bore the brunt of the empire's defensive and deterrent needs. As with the previous cases, this analysis also excludes allied forces for purposes of contingency planning.

The preceding calculations allot 4 pebbles to the United States.[6] Figure 4 represents the strategic schematic of the United States in the post-Cold War world, which has a complexity of 1. It substantially resembles the pentagonal example from Figure 1; but, because of the ability to move from the United States directly to either Europe or the Near East, this figure provides a somewhat easier set of solutions, and fractionally reduced complexity. Rather than recapitulate the analysis of the pentagonal example, the most

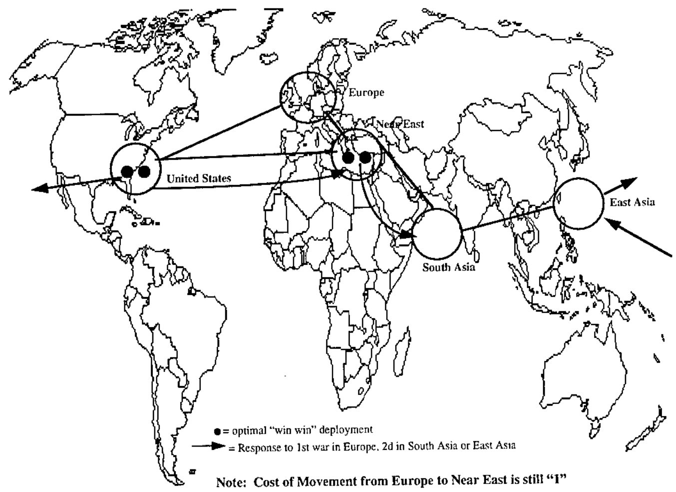


Figure 4. Strategic Schematic of The United States

relevant strategic options and costs can be easily displayed in Table 2.

For purposes of achieving a single win, strategy VII provides the least costly option, and also allows for rapid (one move) response to either of the two unoccupied nodes, Europe and the Near East. The implication for policy here is to de-emphasize Europe and the Mediterranean.[7] However, among multi-nodal configurations, only strategies II and XI provide a win-win option; and this approach creates, in the former instance, a powerful presence in the Levant whose implications, for the Balkans, the Maghreb and the rest of the Arab world, would

Table 2. U.S. Strategic Options

	Node					
Strategy	United States	Europe	Levant	South Asia	Far East	Cost, 1 win
I	2	2				4
II	2		2			3
III	2	1	1			4
IV	1	1	2			3
V	1	2	1			4
VI	1			1	2	4
VII	2			1	1	2
VIII	4					6
IX	3	1				4
X	1	3				4
XI	3				1	4
XII	3				1	5

Note: Only Strategies II, VIII and XI achieve "win win" conditions.

be profound. In the latter case, a much greater standing presence in the Muslim world is also implied. As both strategies call for increased presence in those areas where political pressures for "minimized presence" are greatest, they might prove difficult to pursue in practice.

In terms of deterrence, Strategy II would respond at a cost of 1 to any first war; but deterrence of a second conflict could be weakened, as two moves might be needed to respond in the latter case. For example, with 2 pebbles each in the United States and the Eastern Mediterranean, a war in Europe would necessitate moving 1 from either node. If forces came from the U.S., then responding to a second war in the Far East would require staging forces from the Eastern Mediterranean through the United States on their way to this war, at a cost of 2. If the initial response came from the forces in the Mediterranean, then a second war in South Asia would necessitate a response from the United States, through the Mediterranean, then on to the scene of the conflict.

It should also be noted that pulling U.S. forces home (VIII) raises response time substantially, to treble the cost of the most effective single-war strategy (VII), and double that of the only multi-nodal strategy that allows a "win win." However, an "all-U.S." configuration does allow a win-win situation, albeit at costs that could run as high as 3 to reach the second war if, for example, the first war were in Europe, and the second broke out in South Asia. This implies an even greater weakening of deterrence through adherence to a strict depth defense that would allow the initial overrun of peripheral areas.

Finally, if, for whatever reason, the United States were to shift to a one-war strategy, then it could achieve a win with only three pebbles, a 25% reduction *beyond* those cuts currently envisioned to reach base-force goals. This would, however, put the United States in a position from which it could not, given its resources, cope with a second conflict in a timely manner. Also, response to distant crises would be slower, weakening deterrence.

With regard to the insights provided by the historical examples, the key point is that both cases confirm the logical correctness of shifting to depth defenses under conditions of scarcity. However, the reduction in forward presence, coupled with the time needed to respond to distant crises, also showed that depth defense may vitiate deterrence stability. Thus, there exists a tension between efficient defense, employing an economy of force, and robust deterrence. Of course, this finding rests on the notion that successful deterrence generally depends upon the maintenance of a viable defensive force (as opposed to "flag-waving" detachments), either in the threatened region, or near enough so as to provide a rapid response. This formulation has received strong theoretical and empirical support in recent years, notably from Mearsheimer (1983), Huth and Russett (1984, 1988) and Huth (1988). Others (Jervis, Lebow and Stein, 1985; and Payne, 1992) have pointed out the consistent perceptual problems that aggressors have exhibited throughout history, reinforcing the point that robust deterrence may depend on having viable defensive capabilities in place, or nearby, in a crisis.

Finally, this formal analysis of the American strategic situation suggests the possibility of shifting deployments in Europe away from the central plains of Germany to the Mediterranean. Such a move improves response time to any South Asian crisis. It would also, no doubt, shore up deterrence against the spread of unrest from the Balkans, or against any sort of missile threats from the Maghreb. One policy implication might thus be to give very serious consideration to Albania's application for NATO membership. Given the apparent waning of the Russian threat, a southward shift of U.S. forces, strategically optimal from a graphing perspective, could take advantage of an opportunity to enhance security throughout the Mediterranean, from the Balearics to the Balkans.

CONCLUSION

Several insights have emerged from applying the graphing methodology described herein. Most significantly, this approach helps both in identifying optimal strategies and determining the minimum level of resources required for effective deterrence and defense. Also, the methodology can detect and then quantify the effects that reduced or increased force levels

have on marginal costs. Finally, this study provides a verdict on the prospects for successful "win win" depth defense, confirming that the "base force minus" reductions which the U.S. military currently endures do not, of themselves, compromise its ability to fight and win two regional conflicts with near simultaneity. Further, if American strategy were to shift to a willingness to fight one war at a time, U.S. forces could reduce to even lower levels. In this case, though, deterrence could be substantially weakened, as the response time needed to reach most crises would lengthen considerably.

One should remain cautious regarding the applicability of formally- derived findings such as those generated in this study. While the mathematics of graphing may imply an ability to reduce forces sharply, thanks to optimized redeployments, the political reality of maintaining minimal standing forces is that considerable institutional opposition will likely arise. Indeed, one of the lessons of the Vietnam War for the American military was that overwhelming force should be applied whenever possible. Examples of this philosophy abound, from the massive expeditionary forces that descended upon one Cuban construction detachment in Grenada (as opposed to the losing "shoestring" forces that invaded Cuba in 1961 at the Bay of Pigs), to the preponderant forces that quickly overran the various militia-like "dignity battalions" of Noriega in Panama. Even the Iraqis, the strongest of America's recent adversaries in war, found themselves confronted by overwhelming air, sea and ground forces of the United States.[8]

The point here is that there may be organizational, bureaucratic political, or even prudent strategic reasons for wanting to have more forces than required by some theoretically-derived minimum level. After all, moving toward any minimum requirement entails risk. One can hardly quarrel with such "perturbing" factors, in terms of their existence. However, one may employ the arguments advanced in this study to determine "bottom line" requirements accurately. In this regard, the graphing methodology provides a value-neutral framework for analysis that may help to mitigate the natural tensions that arise between a desire for robust security and the need to operate under ever more constrained fiscal circumstances.

Finally, this methodology has generated the insight that marginal increases over bare-minimum force requirements can have very substantial effects in terms of improving coverage of key areas and response time in an international crisis. In the generic model (the "pentagon" of Figure 1), one pebble more than the minimum reduced costs of response by 50%. This means, in theory, that deterrence stability could be decisively enhanced with modestly larger forces. The implication for American defense policy may be that, even though the base force can achieve the "win win" grand strategy, its minimalist nature may weaken deterrence in crisis, leading to the outbreak of, and American involvement in, more wars. If this is true, then it behooves U.S. policymakers to contemplate the deterrent and stabilizing effects of having just one more "pebble."

ENDNOTES

1 Friedberg (1988), in his examination of Britain in the pre-World War I period, notes both some efforts to retrench, but also reflects on the terrible tension caused by "trying to continue to play the part of a world power without being willing to pay for the privilege" (p. 303).

2 For the purposes of this study, it is assumed that all nodes, or regions, are "created equal." That is, one pebble can win anywhere, even though the cost of moving one a great distance may be great, once there it performs as well as it would nearer to home. Much evidence supports the notion that the Roman legions fought at a relatively constant level of military effectiveness across a wide variety of regions. Similarly, the U.S. forces, that had been developed and trained for combat in Central Europe made a smooth transition to desert warfare in the war against Saddam Hussein.

3 The cost structure employed herein also captures the essence of the problem associated with the movement of expeditionary forces: the ability to move depends upon having rearward infrastructure. In U.S. practice, this notion of

the "rotation base" forms a central element in American force projection capabilities.

[4] Similarly, while one may have allies, it is prudent (and a standard practice) in contingency planning to set military requirements on the basis of having to fight on occasion without militarily effective (e.g., the GCC states' armed forces in 1990) allies.

[5] Two other assumptions underlay these calculations. First, it will remain necessary, for the foreseeable future, to move heavy forces by sea to any given region. Second, conventional war has not yet reached a stage at which American information and other high technology, included space-based weapons, have significantly lessened the need for large field forces.

[6] The aggregate requirement for twelve divisions is met by combining the ten Army divisions with the three Marine divisions that will continue in existence. This leaves a slight overage.

[7] It is assumed, for political reasons, that at least 1 pebble will have to remain in the continental United States and its environs.

[8] Indeed, Aspin (1992, 29-34) argues forcefully that the Iraqis had considerably fewer troops in and around Kuwait than the half-million figure commonly assumed. The actual Iraqi order of battle could have had as few as 183,000 troops. Of course, U.S. air and naval mastery made the odds against the outnumbered Iraqis much worse.

REFERENCES

Aspin, Les. 1992. *Defense for a New Era.* Washington, D.C.: GPO.

Aspin, Les. 1993. *The Bottom-Up Review.* Washington, D.C.: GPO.

Boulding, Kenneth E. 1962. *Conflict and Defense.* New York: Harper.

Delbruck, Hans. [1921]1990. *History of the Art of War, Vol. 2, The Barbarian Invasions.* Trans. by Walter Renfroe. Lincoln: University of Nebraska Press.

Doran, Charles F. 1971. *The Politics of Assimilation.* Baltimore: The Johns Hopkins University Press.

Elliott, J.H. 1991. "Managing Decline: Olivares and the Grand Strategy of Imperial Spain." In Paul Kennedy, ed. *Grand Strategies in War and Peace.* New Haven: Yale University Press.

Ferrill, Arther. 1986. *The Fall of the Roman Empire.* London: Thames and Hudson.

Ferrill, Arther. 1991. "The Grand Strategy of the Roman Empire," in Paul Kennedy, ed., *Grand Strategies in War and Peace.* New Haven: Yale University Press.

Fisher, Admiral of the Fleet Lord. 1919. *Records.* London: Hodder and Stoughton.

Friedberg, Aaron L. 1988. *The Weary Titan: Britain and the Experience of Relative Decline.* Princeton: Princeton University Press.

Gibbon, Edward. [1776]1909. *The Decline and Fall of the Roman Empire.* London: J.B. Bury.

Gilpin, Robert. 1981. *War and Change in World Politics.* Cambridge: Cambridge University Press.

Goldstein, Joshua. 1988. *Long Cycles.* New Haven: Yale University Press.

Gupta, Raj. 1993. *Defense Positioning and Geometry.* Washington, D.C.: The Brookings Institution.

Hale, W.K. 1980. "Frequency Assignment: Theory and Applications." *Proceedings of the IEEE*, 68:1497-1514.

Harrison, J.L. 1962. "The Distribution of Feeding Habits among Animals in a Tropical Rainforest." *Journal of Animal Ecology*, 31:53-63.

Huth, Paul. 1988. *Extended Deterrence and the Prevention of War.* New Haven: Yale University Press.

Huth, Paul and Bruce Russett. 1984. "What Makes Deterrence Work?" *World Politics* 36:496-526.

Huth, Paul and Bruce Russett. 1988. "Deterrence Failure and Crisis Escalation." *International Studies Quarterly* 32:29-46.

Jervis, Robert, Richard Ned Lebow and Janice Gross Stein. 1985. *Psychology and Deterrence.* Baltimore: The Johns Hopkins University Press.

Jones, Arnold. 1964. *The Later Roman Empire (284-602): A Social and Economic Survey.* 3 vols. Oxford: Basil Blackwell.

Kemeny, J.G.and J.L. Snell. [1962]1972. *Mathematical Models in the Social Sciences.* Cambridge: MIT Press.

Kemp, P.K. 1964. *The Papers of Admiral Sir John Fisher.* London: The Navy Records Society.

Kennedy, Paul. 1976. *The Rise and Fall of British Naval Mastery.* London: Ashfield Press.

Kennedy, Paul. 1987. *The Rise and Fall of the Great Powers.* New York: Random House.

Luttwak, Edward. 1976. *The Grand Strategy of the Roman Empire.* Baltimore: The Johns Hopkins University Press.

Marder, Arthur. 1940. *The Anatomy of British Sea Power: A History of British Naval Policy in the Pre-Dreadnought Era,* 1880-1905. New York: Alfred A. Knopf.

Macmullen, Ramsay. 1976. *Roman Government's Response to Crisis, AD 235-337.* New Haven: Yale University Press.

Mearsheimer, John. 1983. *Conventional Deterrence.* Ithaca: Cornell University Press.

Modelski, George. 1987. *Long Cycles in World Politics.* Seattle: University of Washington Press.

Modelski, George and William R. Thompson. 1988. *Seapower in Global Politics, 1494-1993.* Seattle: University of Washington Press.

Organski, A.F.K. and Jacek Kugler. 1980. *The War Ledger.* Chicago: University of Chicago Press.

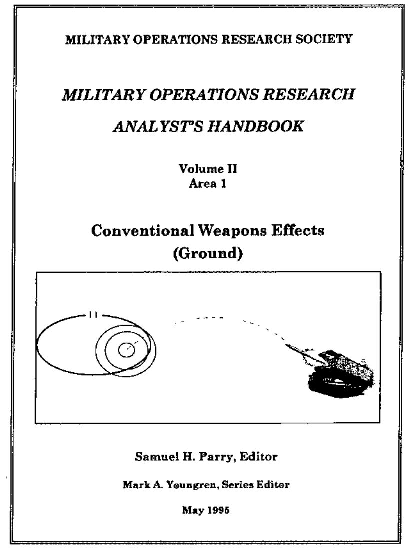
Owens, William A. 1995. *High Seas: The Naval Passage to an Uncharted World.* Annapolis: Naval Institute Press.

Payne, Keith B. 1992. *Deterrence and the Lessons of History.* Washington, DC: National Institute for Public Policy.

Tucker, A.C. 1973. "Perfect Graphs and an application to Optimizing Municipal Services." *SIAM Review,* 15:585-590.

ANALYST'S HANDBOOK SUPPLEMENT NOW AVAILABLE

Area I of Volume II, *Conventional Weapons Effects*, is now available through the MORS office. It is intended to be part of the initial notebook in the *Handbook* series, along with Volume I (*Terrain, Unit Movement, and Environment*) and Volume II, Area II (*Search, Detection and Tracking*).



The *Conventional Weapons Effects* (*Ground*) area was edited by Sam Parry, who has been educating and advising Army OR analysts for many years. It is designed to provide a quick reference for models that represent the effects of a conventional attack against ground targets. It emphasizes the ground battle; publication of algorithms for naval and air engagements is planned for the future. It cannot, of course, include every algorithm used by military OR practitioners, but it is an attempt to characterize those that are most commonly used.

The cost of Volume II Area I is $8.00. As always, Volume I, which includes a three-ring binder for future supplements, is available for $25.00. **Combine both and save: $30.00.** All prices includes freight. Please call the MORS Office for international shipping rates.

✂- -

Volume II Area I: _______ Volume I: _______

Volume I and Volume II, Area I: _______

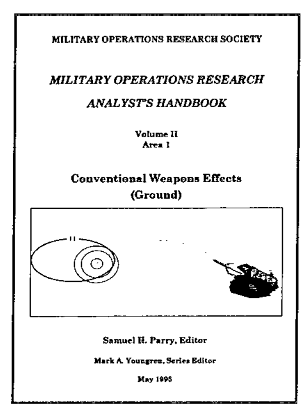
Total Enclosed: _______

Name:____

Shipping Address:_______________________________________

____

____



City:____________________ State:__________ Zip Code:__________

Make checks payable to MORS

Military Operations Research Society ● 101 S. Whiting Street ● Suite 202 ● Alexandria VA 22304 ● (703) 751-7290

The Need to Represent a Wide Variety of Battle Types in Air-Ground Combat Models

Dr. Patrick D. Allen

Cubic Applications, Inc.
4550 Third Ave., SE
Lacey, WA 98503

ABSTRACT

Most existing combat models are severely limited in the types of battles represented in a given scenario. Only a few types of battles are included in most aggregate models, and more detailed models do not adequately represent the aggregate effects of the majority of real-world types of battles. As a result, most combat models emphasize the benefits of lethality and attrition while providing a poor representation of maneuver effects and the benefits provided by C3I systems. Suggestions are provided on how to incorporate a wider variety of battle types into existing and future aggregate and detailed air-ground combat models.

INTRODUCTION

Performing analysis or training with an inadequate set of battle types in our combat models is similar to performing analysis or training with an inadequate set of scenarios. For many years, the military analytic community was plagued with the desire to perform all analysis using a single "blessed" scenario. This resulted in what at best could be considered a single-point estimate of the nearly infinite space of possible outcomes. At worst, it represented a force structure or system "optimized" for a single scenario that had an infinitely small probability of occurrence.

Since the break-up of the former Soviet Union, the need for multi-scenario analysis has become more widely accepted, especially since there is significant uncertainty as to the future threat, the likely theater of operations, and size of the U.S. force structure. Just as there has always been a need for multi-scenario analysis, there was and still is an equally strong need to represent a wide variety of battle types within air-ground combat models. This is true both for aggregate combat models and more detailed combat models, although the nature of the problem and the proposed solutions are different for each.

In aggregate combat models, the most commonly represented battles include a Blue attack against a Red defense, a Red attack against a Blue defense, a Red attack against a Blue delay, a meeting engagement, and a static engagement. A few aggregate combat models, such as the Joint Integrated Contingency Model, include different phases of battle (such as breakthrough, exploitation, and pursuit) and different degrees of attacker's preparations. However, few, if any, existing models represent flank attacks, counterattacks, river crossing operations, and combat during passage of lines operations. Although these types of battles occur frequently throughout the history of combat and are often defined as *key* historical battles, they are not well represented in our aggregate combat models. For example, in the Theater-Level Combat (TLC) model (previously under development at RAND), the assets of the flanking force were simply considered another head-on force attacking the defender, so that the flank battle was thereby assessed as just another head-to-head attack. Thus the unique factors associated with these types of battles are often ignored in analysis and training.

Nor are these types of battles adequately represented in more detailed combat models. Although proponents of models such as the Corps Battle Simulation or the Janus model claim that these models adequately represent, for example, a flank attack, even these models do not represent the factors that make such battle types unique. (Any sample model shortcomings mentioned in this article are not intended to imply that the model is any worse than any other model, but are simply used as examples that the state-of-the-art of air-ground combat models require substantial improvement in representing basic, first-order effects.)

Let us say, for example, that a corps is attempting to attack an enemy army on the flank. In reality, because of the limited information available to the defender and the actual time required to respond in a cohesive manner, such a flank attack or counterattack tends to succeed. It takes a significant amount of time for a defending force to first, identify that such an attack is occurring, and second, how to best respond to that threat.

In a model, however, matters do not necessarily proceed as they would in reality. Although ground-truth is not provided to the players on each side, all players in the model share the same perception database. Therefore, whatever one unit knows, all units know. As a result, as soon as one defending maneuver company is engaged by three flanking battalions, all remaining defending units are "alerted" to this flank

threat and immediately respond. Moreover, each player can "micromanage" the units under his command so that the flank threat is immediately neutralized. Few flank attacks ever succeed in this type of training model, since the receiver of the flank attack has too much information and is allowed to react too quickly. This problem of a readily shared perception of reality is the same both for entity-level models, such as Janus, and models that represent aggregate units.

The models described above, however, allow the defender to quickly and unrealistically turn a flank attack into a head-on attack, and a counterattack into an attack against a hasty defense. Because weapon-system level models also allow for an immediate perception of the flank threat once detected by any friendly element, they share the same problems as aggregate models. As a result, both aggregate and detailed combat models tend to treat almost all battle types as head-to-head battles, regardless of how they began.

Ignoring the distinctive nature of flank attacks is one example of how our existing models tend to focus on lethality and attrition, while discounting assets and actions that support maneuver effects. By failing to adequately represent the different types of battles in our models, we are performing a great injustice to nonlethal combat systems. The effects of mobility and countermobility operations, the operational benefits of C3I systems, and the representation of operational art are all sorely lacking in our existing air-ground combat models. For example, since most models do not assess a flank attack any differently than a head-on attack, such a model would ascribe no measurable benefit to a clever C2 decision.

The rest of this article will focus on the features of several types of battle, what makes each unique, and how each can be adequately represented in our existing combat models. A final section comments on the implications of failing to address different battle types in analysis and training. The scope of this article is limited to mid- to high-intensity conventional combat, and does not include operations other than war or integrated warfare.

PHASES OF BATTLE

In the assault phase of battle, a defender with a cohesive defense tends to have the advantage over the attacker. As a result, the loss rate of the attacker tends to be higher than that of the defender, thereby indicating an advantage to the defender. At least, this is how the assault phase is represented in most combat models.

Historically, however, the defender tends to lose more forces than the attacker (McQuie, 1989). This is because if the attack is successful, then the exchange rate shifts in favor of the attacker during the breakthrough, exploitation, and pursuit phases.

Unfortunately, most models represent only the assault phase, thereby always placing the attacker at a disadvantage. Such a representation actually models an infinitely elastic defender who is always cohesive and whose line never ruptures, regardless of how fast he is in retrograde. Unless our models represent all different phases of battle, they will remain biased against the attacker and in favor of the defender (Allen, 1989).

The basic aggregate parameters that change in these different phases of battle are combat intensity, exchange rate, and movement or penetration rates. For example, during a breakthrough, the combat intensity decreases in terms of overall combat losses, but increases in terms of surrender and isolation. At the same time, the exchange rate should switch from being favorable for the defender to favorable for the attacker, and the advance or movement rate of the attacker should increase.

At a more detailed level, the losses of the less-mobile defender's assets are higher once they are cut off and surrounded. In addition, the defender will recover fewer assets for repair, while the attacker, who owns the battlefield after the battle, will recover a much higher percentage. For example, on the Syrian Front in the '73 War, the Syrians lost 1150 tanks, of which the Syrians recovered at most 25 percent (867 were recovered by the Israelis), while the Israelis lost 250 tanks, of which the Israelis recovered and repaired 150, or 60 percent (Herzog, 1982, p. 341).

Moreover, it will be difficult for the defending force to act in a cohesive manner during the

breakthrough, exploitation or pursuit phase of battle. A detailed model should penalize the effectiveness and response times of defending forces suffering a breakthrough, exploitation or pursuit. For the same reason as described above regarding a flank attack or counterattack, a detailed model should not allow the participants too much information.

ATTACKER'S PREPARATIONS

Another combat effect not represented in most models is the preparations of the attacker. Most models give an advantage to the defender for having prepared defenses, but few represent a benefit to the attacker for a prepared attack, or a penalty for a hasty attack. For example, in reality, if an attacking force advances quickly at 80 kilometers in one day and then suddenly leaps into a hasty attack against a prepared defense across a river, the attacker should have a very rough time of it. The artillery support, especially the ammunition, is still relocating forward. The rapidly advancing force is advancing on a narrow front, requiring deployment onto a broader front to seek out weak points in the defender's line. Not enough time is available to reconnoiter, plan, and rehearse the assault. As a result, a hasty attack has little chance of succeeding against a prepared defender. However, our models generally do not distinguish between different types of attack.

To represent the benefits of attacker's preparations (and the penalties of foregoing preparations), several basic parameters need to be changed. In aggregate combat models, the amount of force that may be entered into battle should initially be small due to a rapid advance across a narrow frontage. This frontage, and the amount of force the attacker may bring to bear, will increase over time. Artillery assets should be especially constrained, primarily due to the lack of adequate munitions to perform a prepared attack. Over time, the ammunition stockpile near the guns will grow. In addition, proper reconnaissance, rehearsal, planning and coordination—all of which take time—may increase the attacker's effectiveness and improve the exchange ratio in favor of the attacker. (Allen, 1992)

A similar set of changes are necessary for an adequate representation of attacker's preparations in more detailed models. Although the frontage representation is more accurate, the availability of artillery assets needs modification. In most models, the quantity of artillery rounds located with each gun are not tracked explicitly, but stored in some generic stockpile. As a result, if an artillery tube is within range, it is allowed to fire as long as its side has sufficient ammunition. In reality, if the artillery units have relocated as much as 30 km forward, their ammunition will take more time to reach them. (In one exercise, the attacker relocated his artillery 30 km forward, but there was no reduction in firing rate at the new location. Since the attacker was represented in the model as having a large enough stockpile in the whole force, it was assumed that the rounds were available where needed.) This delay in ammunition delivery can be represented by reducing the firing rate of artillery units during the preparation period, or by more explicitly representing the location of ammunition on the battlefield (which carries its own representation problems in combat models). More detailed models also need improved representation of the effects of mission planning, coordination, and rehearsal, since these are not represented even in highly detailed models.

AIR AND GROUND INTERACTIONS

Several areas require improvement in the category of air and ground interactions. Many combat models overlook the different degrees of effectiveness of air-to-ground and ground-to-air engagements depending on the situation. For example, many models use a constant effectiveness parameter for each type of air defense asset in a unit, regardless of the status or posture of the asset or unit. In reality, however, a unit that is moving will always have some of its air defense assets moving with it. As a result, not all of its air defense assets will be available to engage attacking or passing aircraft. Air defense assets on the move should have a reduced effectiveness in our models, compared to air defense assets prepared to engage targets.

Similarly, many models represent a constant effectiveness for air-to-ground attacks for each

type of aircraft regardless of the posture of the target being attacked. Many models define a fixed number of kills per sortie by type of aircraft regardless of the posture of the attacked unit, and sometimes regardless of the type of unit being attacked. In reality, units in defensive positions or assembly areas are much more difficult to engage than units in the open or on roads. The air-to-ground assessment processes need to be sensitive to the posture and type of the attacked units.

The representation of close air support (CAS) tends to assume that a forward air controller (FAC) is being used to guide the attacking aircraft. This highly effective combination of CAS and FAC is assumed to apply regardless of the number of CAS aircraft attacking a given area per unit of time. Most models allow hundreds of CAS aircraft to engage ground targets as though each were carefully managed by a FAC. In reality, a well-trained FAC can guide only a very limited number of CAS aircraft against observed or known targets per hour. A simple fix would be to reduce the effectiveness of CAS sorties if the number of sorties exceeds what a FAC can handle, or in the absence of a FAC.

Our models also need to reflect the nonlethal effects of air-to-ground attacks. In most aggregate and detailed combat models, the only benefit of an air attack is to kill enemy ground assets. This contributes to making our models attrition-oriented rather than maneuver-oriented. In reality, much of the desired effect of airpower is to delay the movement of enemy ground forces so that the friendly ground forces can engage the enemy in smaller units. The air-land battle doctrine strongly emphasizes this concept, yet our models do not adequately reflect nonlethal effects. It is also important to represent the effects of delay, disruption, and demoralization caused by air-to-ground attacks in both aggregate and detailed models.

FLANK ATTACKS AND COUNTERATTACKS

These two types of battles are discussed together because the problems associated with their representation are related. For these types of battle, we will distinguish between the representations in aggregate low-resolution models and in detailed high-resolution models.

In aggregate models, a force multiplier is used to reflect the benefit to the attacker of a flank attack. These multipliers tend not to exceed a factor of three, thereby causing most simulated counterattacks to fail. In reality, a flank counterattack may be able to roll up the flanks of a much larger force, as was demonstrated repeatedly by Israeli forces in the '67 and '73 wars. (Herzog, 1982, pp. 168-291)

This need not be represented by a large force multiplier, but could be represented by a series of sub-battles in which each engagement is favorable to the flank-counterattacking force. This approach better reflects the mechanism by which flank attacks and counterattacks succeed: by engaging pieces of the enemy under conditions favorable to the counterattacking force, as quickly as possible, before the defender can prepare a cohesive defense. This was the principle used by the German Army in World War II (DePuy, 1980). Unfortunately, many models currently allow the counterattacked force to shift quickly into at least a hasty defense, often causing the counterattack to be assessed as a failure.

In high-resolution models, the problem stems from the availability of too much information to each side and the ability to micromanage forces. Since intelligence models tend to provide a single perception database to all players, the players know almost immediately when and where the counterattack is occurring. Even when intelligence is limited, players are allowed to micromanage component units to quickly achieve a cohesive reaction. As a result, a cohesive defense is prepared much too quickly, when, in reality, the counterattacked force should have been virtually destroyed. Due to these shortcomings in combat simulations, counterattacks against larger, flanked units do not appear to be worthwhile operations.

A recommended solution is to degrade the information available, delay the information over time, and delay the reaction times of the threatened forces. All of these measures would contribute to a more realistic representation of flank attacks and counterattacks.

RIVER CROSSING OPERATIONS

Some models have difficulty representing certain types of assault, such as river crossing operations. In most river crossing operations, part of the force secures the opposite side of the river. This far-shore force is usually infantry-heavy, with limited antitank and antiair capability and no defensive preparations. This makes the force particularly vulnerable to artillery fire and armored counterattack. Helicopters and fixed wing aircraft can help protect the far-shore force, but tend to be vulnerable to air defense fire. In addition, in some models, engineer bridging assets are often of no value in river crossing operations. Few models account for all of these unique features in a river crossing operation.

Aggregate combat models need to represent the vulnerabilities of the bridgehead force to a combined arms counterattack. Similarly, they should include the need for an adequate number of bridging or ferrying assets, the time required to put each into operation, and the reduced rate at which crossing forces can be placed in the bridgehead.

More detailed combat models need to separate the crossing elements of a force (e.g., infantry, light assets, floatable assets) from the rest of the force for combat assessment purposes. If the model accounts for combined arms effects, then the bridgehead will be at a disadvantage if it is counterattacked by a combined arms force.

Even when detailed combat models do account for different types of bridging units, they do not tend to account for the consumption of bridging equipment. In most detailed models, the bridging capability is always considered present with the unit, regardless of the number of bridges already constructed by that unit. When a bridge is placed at a given location, however, it is usually left in place for follow-on forces and logistics. Each engineer unit can undertake only a limited number of bridging operations. Our models should not overestimate the number of bridges that can be constructed over time.

PASSAGE-OF-LINES OPERATIONS

Even though passage-of-lines is an operation rather than a type of battle, a problem with its representation occurs in models that assess no penalty for being engaged during such operations. During a passage-of-lines operation, whether advancing or withdrawing, both the passing and the standing force are more vulnerable to enemy action than either would be if only one of the forces were present. If combat were to occur, both units would have degraded combat capability, so that the two units in transition would not be as effective as either unit in place.

Most models ignore this fact, however, and allow forces to be swapped on and off line and passed with no penalty. Worse, if the density of forces has increased in the combat sector of the model, then the passing side actually has a combat advantage in combat algorithms, rather than a disadvantage as would occur in reality. Because the models do not account for the awkward posture of passing units for combat, they present an unrealistic capability to swap fresh units on and off line under enemy pressure.

Aggregate combat models need to make a passage-of-lines operation explicit, and should include penalties for being in combat during a passage-of-lines operation. Rather than benefiting from the increased number of assets in the combat area, units should be penalized for being overcrowded, disorganized, and mixed in with other friendly units.

More detailed combat models also need to penalize a unit attempting a passage-of-lines operation, rather than allowing it to benefit from having more assets available. For example, most detailed combat models do not penalize a unit for having too many assets in the same location. In entity-level models, such as Janus, several tanks can be stacked in exactly the same location without penalty. As long as all assets have line of sight to the target, they can all fire, regardless of how much they would interfere with each other in reality. Some representation of direct fire fratricide, which is not represented in most high-resolution combat models, would help.

OPERATIONAL ART

To improve the representation of combat operations in our models, we also need to represent many of the first-order effects of the operational arts. The key tradeoffs in operational art

are forces, space, and time. The rate at which these three components interact as a function of the echelon is the key to understanding and representing operational art. However, our models severely lack adequate representation of these tradeoffs.

The most important factor is the time required to coordinate the activities of component units into a cohesive force. Doctrinally, a division takes about 24 hours to initiate a new action, a corps about 48 hours, and an army 96 hours. There is significant inertia behind a higher-echelon force that has been committed to an operation, but most models allow forces to change plans and implement new plans at a moment's notice. In aggregate models, formations are moved and actions initiated much more quickly than could be accomplished in reality.

In detailed models, small formations are micromanaged so that large formations act at nearly the same speed as the smallest unit represented in the model. For example, if the model resolves units down to company size, players are allowed to move each company in a division individually. If the model represents company movement rates of about 25 km/hour, then the composite division moves at about 20 km/hour sustained over a day. In reality, actual divisional daily movement rates tend to average less than 10 km/hour, including even the most rapid advances during WWII and Operation Desert Storm. (Vigor, 1983, p. 112; Savkin, 1965, pp. 4-5; Adler, 1991, p. 101.)

The ability to micromanage smaller units in a model so that they act more efficiently than the larger unit they comprise also affects combat assessment, such as the flank attack example described above. Entity-level models also suffer from groups of entities acting at speeds in excess of how fast we observe such groups acting in reality.

The purpose of operational art is to balance forces, space, and time to achieve the objectives while fighting in as favorable a situation as possible. The lack of representation of many basic aspects of operational art precludes our models from adequately training commanders and staff in the tradeoffs associated with operational art, or from performing proper analysis on systems and procedures that can assist operational art. This is because, even with an intelligent model, too much information is available to the players, who can coordinate the actions of many small forces without representing the delays that would occur in planning and executing complex operations.

OTHER AREAS

Because of the increased emphasis on the use of simulations in training and analysis, there is an increased need to make sure our simulations are up to the task. Of increasing importance are what have been traditionally considered "soft" factors, such as C3I, information warfare, psychological operations, civil affairs, political issues, environmental issues, economic issues, and the representation of friction in all operations. For example, as the U.S. force structure decreases in size, the value of C3I to prepare forces to be in the right place at the right time should be reflected in our models.

In addition, good C3I reduces the level of uncertainty faced by the units, thereby reducing the likelihood of fratricide. Models that attempt to represent fratricide without some measure of uncertainty do not actually address the underlying cause of fratricide. A model that defines a specific type of battle as always incurring a specified amount of fratricide misses the reason that fratricide occurs in the first place. If the model does not vary the fratricide rate as a function of the degree of uncertainty faced by friendly forces, it does not actually represent fratricide.

OVERALL EFFECT

If simulations do not represent the benefits of different types of battles, then no payoff is associated with employing good command and control. For example, if there is no significant benefit to a flank attack, then one is left with only head-on assault options. Many current combat models tend to ignore the benefits of maneuver and emphasize attrition effects. These limitations also affect the representation of operational art.

The limitations in current combat models need to be addressed to better support U.S. military training and analytic requirements. Although many of these problems are pervasive

across almost all simulations, much of the user community is not aware that these problems exist. Because of the complexity of many models, it can be difficult to understand how a detailed model might not adequately represent aggregate effects. Therefore, many model proponents assume that their models account for effects that they do not actually represent.

Military officers understand how important these factors can be, but many do not realize that current simulations do not already account for these fundamental effects. When representatives of the user community focus on tactical-level issues, it is difficult for problems associated with operational-level issues to be identified and corrected. When a given subject is not understood at one level of resolution, a commonly attempted solution is to add more detail. Unfortunately, this increase in detail actually tends to obscure the primary issue.

A serious difficulty with current combat models is that the types of problems that should be considered first-order effects, such as an adequate representation of combined arms effects, types of battles, friction, and passage-of-lines operations, are not usually raised to the top of the list by the user community. The main reason appears to be that the user community simply assumes these first-order aggregate effects have been adequately represented in the models, so that increasing the details of the models becomes the focus of attention. Frequently, only model improvements related to increasing detail are high enough on the priority list to be funded, precluding the improvement of our models in the area of realistic aggregate effects. It is hoped that this article will raise the issues of adequate representation of these factors to both the model builders and the model users. Moreover, we hope that these issues can be raised in priority so that they can be addressed, rather than ignored.

REFERENCES

Adler, Jerry, "The Final Push," *Newsweek Commemorative Edition on America at War*, Spring/Summer, 1991.

Allen, Patrick D., *Situational Force Scoring: Accounting for Combined Arms Effects in Aggregate Combat Models*, RAND, N-3423-NA, 1992.

Allen, Patrick D., "Response to 'Numbers in Combat,'" *PHALANX*, September, 1989.

DePuy, William, *Generals Balck and Von Mellenthin on Tactics: Implications for NATO Military Doctrine*, the BDM Corporation, BDM/W-81-077-TR, DNA001-78-C-0114, December, 1980.

Herzog, Chaim, *The Arab-Israeli Wars*, Random House, New York, 1982.

McQuie, Robert, "Numbers from Combat," *PHALANX*, June 1989.

Savkin, Col. V. Ye., *Rates of Advance*, Military Publishing House of the Ministry of Defense, USSR, Moscow, 1965.

Vigor, Peter H., *Soviet Blitzkrieg Theory*, St. Martin's Press, New York, 1983.

Do YOU think you know what the future holds for us...

"Computers in the future may weigh no more that 1.5 tons."
—*Popular Mechanics*, forecasting the relentless march of science, 1949

"I think there is a world market for maybe five computers."
—Thomas Watson, Chairman of IBM, 1943

"But what ... is it good for?"—Engineer at the Advanced Computing System Division of IBM, 1968, commenting on the microchip.

"There is no reason anyone would want a computer in their home."
—Ken Olson, President, Chairman and founder of Digital Equipment Corp., 1977

"The wireless music box has no imaginable commercial value. Who would pay for a message sent to nobody in particular?"
—David Sarnoff's associates in response to his urging for investment in the radio in the 1920s.

"The concept is interesting and well-formed, but in order to earn better than a C, the idea must be feasible."—A Yale University management professor in response to Fred Smith's paper proposing overnight delivery service. (Smith went on to found Federal Express Corp.)

"Heavier-than-air flying machines are impossible."—Lord Kelvin, President, Royal Society, 1895.

"Professor Goddard does not know the relation between action and reaction and the need to have something better than a vacuum against which to react. He seems to lack the basic knowledge ladled out daily in high schools."—1921 *New York Times* editorial about Robert Goddard's revolutionary rocket work.

"640K ought to be enough for anybody."—Bill Gates, 1981

Request for Application

64th MORSS Symposium — 18 - 20 June 1996
Theme: *Leveraging Technology for the Military Analyst*
Fort Leavenworth, KS

Please type (or print **legibly**) the information required on each line below. Use your **business** address. Mail to: MORS, 101 S. Whiting St., #202, Alexandria, VA 22304 or FAX (703) 751-8171 your request to the MORS office. Questions, call the MORS office at 703-751-7290. An application, security voucher, preliminary program and instructions will be mailed in late January 1996.

Name (Title, Rank and Service) ______________________________
(Active duty applicants must include their rank)

Name you want on your badge ______________________________

Social Security Number ______________________________

Company, Command or Agency ______________________________

Business Address or PO Box ______________________________

OR Home Address ______________________________

City, State and Zip Code ______________________________

Phone ______________ DSN ______________ FAX ______________

Email ______________________________

❑ **Please check here if you require special assistance.**

If you are interested in presenting a paper at the 64th MORS Symposium, you may contact the Working Group Chair directly or mail/FAX this form along with your abstract to the MORS Office. The Working Group Chairs addresses and phone numbers are listed on page 2.

❑ YES, I am interested in presenting a paper in Working Group(s):

WG ________ WG ________ WG ________ WG ________

❑ Attached is an unclassified abstract which has been *Approved for Public Release*. (Please follow directions listed under **Procedures for Submitting Papers**.)

❑ I would like to help with your working group.

Two Effects of Firepower: Attrition and Suppression

Wayne P. Hughes, Jr., FS
Senior Lecturer
Department of Operations Research
Naval Postgraduate School
Monterey, California 93943

ABSTRACT

A new basis for the quantitative study of ground combat is introduced that argues the inadequacy of attrition models and the need to incorporate the effects of suppression of the enemy with firepower. A quantitative approach to suppression of enemy fire is offered. Then an analysis shows that the effect of own fire in suppressing enemy fire will, in suitable, frequent circumstances, reverse the conclusions derived from the Lanchester square law, so that the squared term determining the victor is unit firepower instead of the numbers of units engaged.

A fundamental aim of physical science is to describe its processes with dynamic models, mathematical if possible. The same aim, often implicit, is true of descriptions of combat phenomena in military operations research. Models of military operations are necessarily more abstract and approximate than those of physical science. This is especially so because the scientific study of combat operations is complicated by human presence. The problem, of course, is to reduce the enormous effects of "human factors," to an understandable construct, or paradigm. An excellent statement of the problem is in Davis and Blumenthal's monograph, *The Base-of-Sand Problem: A White Paper on the State of Military Combat Modeling* [1991].

For more than a decade The Military Conflict Institute (TMCI) has been addressing the problem via a theory of combat that derives from six axioms. A goal of TMCI is to express the theory in a study, thus far unpublished.[1] As something of a status report, I undertook a Naval Postgraduate School research paper to digest what seemed to be the essence of TMCI's work. It is entitled "Combat Science: An Organizing Study" [1993]. The present paper is an exposition of a small but important consequence of the two endeavors.

We may come at our subject with the following question which arises from the remarkable results of the Gulf War:

> What analytical proposition helps to explain the evidence that much superior combat power when properly applied will result in disproportionately small losses to the winner while he achieves his objective?

The relevant principles from Combat Science are:

- Military force, or combat power, is a real phenomenon, the results of which are observed by its effects on the enemy in battle.
- The observable effects of combat power are not merely physical (producing casualties) but also mental (persuading the enemy of our superiority) and spiritual (diminishing enemy morale and will to continue fighting.)

For purposes of this paper, we will look carefully only at the most measurable manifestation of combat power's mental effect, which is suppression of enemy actions, specifically his return of fire. Its spiritual effect to diminish enemy morale plays no part in the computations that follow, but may be seen to be an unquantified bonus.

Let us begin with a fresh look at the Lanchester square law from the perspective of combat science. As it was conceived and employed by Chase [1902], Fiske [1905], Osipov [1915] and Lanchester [1915] the square law describes combat as a purely physical phenomenon:

[Unit fire's physical effect in casualties imposed/minute]

x

[# of physical units firing]

=

[Fewer enemy units]

Starting with Osipov, who reached his own unique conclusion as to the appropriate relationships, there have been many objections that the classical square law does not fit the historical battle casualty data for ground combat. The most commonly cited reason is (properly enough) that the law cannot apply when the required conditions are not present in the battle. The conditions are that each participant must be able to fire at each participant on the opposing side; and incapacitated opponents must be known at once, so that fire is distributed only against active opponents. Patently these conditions are seldom fully met, with not-so-obvious effects on the square law's applicability.

There are probably three additional major reasons, each related to the other two.

One is insufficient attention to defender advantage. Unit fire effectiveness will normally be greater on the defender side until

the attack is nearly consumated. This advantage comes from the defender's superior posture. When identical unit effectiveness is (wrongly) assumed for attacker and defender the results will appear to show a less-than-square-law effect for the attacker's numerical advantage.

The second reason is that a battle is usually episodic, such that different elements of a force play predominant roles at different phases of the battle. A battle might involve disproportionate losses upon the attacker during his assault and disproportionate losses upon the defender during his attempt to disengage. Different weapons produce casualties more and less effectively during different phases. Historical battle data rarely specify casualty production in this way, and so the episodic effects are disregarded out of necessity.

The third reason is shortage of data with which to measure the effect of suppression of enemy actions from firepower.[2] Since suppression produces no casualties and is a transitory phenomenon that disappears when the battle is over, the effect of its presence is overlooked (or is merely implicit) in almost all combat models, including high resolution simulations. In Chapter 18 of *Understanding War,* T. Dupuy began his fine discussion of suppression with: "There is probably no obscurity of combat requiring clarification and understanding more urgently than that of suppression" [1987].

The heart of the problem is not a refusal to acknowledge the importance of suppression but the lack of an analytical approach that describes the phenomenon and its importance. We wish here to develop a quantifiable model of it. Then we will fulfill the purpose of this paper, which is to show how an appreciation of the cognitive influence of fire can reverse present conceptions of the value of numbers relative to the unit firepower of a fighting force.

We begin with the common form of the Lanchester square law:

$$\frac{dA}{dt} = -\beta B(t) \qquad \frac{dB}{dt} = -\alpha A(t) \quad (1)$$

where α and β are the *constant* unit effectiveness coefficients for A and B respectively, measured in kills per shooter per unit of time. A(0) and B(0) are the initial force strength; A(t) and B(t) are the forces remaining at any later time t; and dA/dt and dB/dt are the rate of losses of A and B respectively at time t.

A solution to equations [1] for end time T is the state equation:

$$\alpha\,[A(0)^2 - A(T)^2] = \beta\,[B(0)^2 - B(T)^2] \qquad (2)$$

Equation [2] shows the well known square law phenomenon, in which A will win a fight when the weaker side is annihilated:

$$\alpha\, A(0)^2 > \beta\, B(0)^2$$

and vice-versa. The classic conclusion from the square law is that the "fighting strength" (Lanchester's term) of a force increases in direct proportion to unit effectiveness and in proportion to the square of the number of forces engaged.

Notably, the casualty-producing effect of fire is regarded as a constant. There is ample evidence that unit effectiveness, α or β, is not even approximately constant but is highly variable. This is true when one is using casualty data taken battle after battle or day after day during a campaign: see for example, D. Hartley [1990, pp. 3-4; and 1991, p. 3]. It is true when one is using casualty data taken from different places on a battlefield: see G. Kuhn [1989], for example. It is true during different phases of a battle: see R. Helmbold [1979], for example.[3] Even though data by phase is seldom available, the variability of fire effectiveness is readily appreciated by reflecting upon a battle's sequence: for example, an artillery preparation phase, an armored assault phase, and an exploitation phase after an enemy position is overrun.

If α (t) and β (t) are also variables, then the Lanchester square law equations must be written:

$$\frac{dA}{dt} = -\beta(t)\,B(t) \quad \frac{dB}{dt} = -\alpha(t)\,A(t) \qquad (3)$$

Two important reasons that α (t) and β (t) vary are: first, a change in range between the fighting units as the battle unfolds; and second, the suppressive effect of each side's fire on the other's fire. For two forces fixed in place with no change in range, the suppression suffered by A will be in consequence of the volume and

accuracy of fire by B, and vice versa. Let us postulate a constant of proportionality, g, such that $g\beta(t)B(t)$ is the time-rate at which each A-shooter's fire is curtailed (measured in volume or accuracy or both). Similarly define a value, h, such that $h\alpha(t)A(t)$ is the rate that each B-shooter's effectiveness is diminished by A's fire.[4] We write:

$$\frac{d\alpha}{dt} = -g\beta(t)B(t) \qquad \frac{d\beta}{dt} = -h\alpha(t)A(t) \qquad (4)$$

The total effect of firepower represented by equations [3] and [4] is twofold. Fire will reduce both the numbers of enemy forces remaining and the lethal effectiveness of the remaining forces. Old equations [3] and new equations [4], however, have dimensions incompatible with each other. After reconciling the dimensions (to killing power lost per unit of time), it is possible to retain all quantities and proceed to explore the relationships between the four equations. There remains, of course, the serious problem of finding data, even rough data, for inputs.

DEVELOPMENT

Let us proceed, now, to develop our principal point. As remarked above, casualties are a permanent consequence of fire, but suppression's effect is only transitory. We usually say combat forces are "pinned down" under fire, but not permanently affected. Arguably the changes in α and β will be much greater during a battle than the changes in force size caused by attrition. In a typical land battle, casualties will be less than 10%.[5] See among many examples, R. McQuie, "Battle Outcomes: Casualty Rates as a Measure of Defeat" [1987].

As McQuie says, something other than casualties must cause the losing side to break off action in a typical battle. In many, perhaps most, battles mission success is obtained by the cumulative cognitive effects of fire on the enemy. The losing side's forces curtail their maneuvers and fire to avoid destruction. The losing side is also increasingly dispirited, with further loss of fighting effectiveness. Or the losing commander thinks that the situation can only go from bad to worse and withdraws to fight afresh from what he hopes will be a more favorable posture.

We may easily explore circumstances in which the volume and precision of fire have a much greater effect to suppress enemy fire than to impose casualties. Define b(t) as the rate of fire in shots by each unit B, and a(t) as the rate of fire by each A. Define s as the rate of reduction of fire imposed on the other side per shot fired, such that:

$$\frac{da}{dt} = -s\,B\,b(t) \qquad \frac{db}{dt} = -s\,A\,a(t) \qquad (5)$$

With equations [5] we measure combat power wholly by its effect on the volume and accuracy of the enemy's return fire during the course of the battle. The model, in its stylized purity, says that a battle is won by the side that sustains its own fire and suppresses that of the enemy. Casualties are an unmeasured byproduct.

The state equation for [5] is entirely analogous to equation [2]:

$$A[a(0)^2 - a(T)^2] = B[b(0)^2 - b(T)^2] \qquad (6)$$

Equation [6] leads to a radically new "law" of combat in which the quantity of fire (measured by its suppressive effect) is the squared term. Now side A will win a battle in which the enemy's capacity to fire back is "annihilated" (fully suppressed) if:

$$A\,a(0)^2 > B\,b(0)^2$$

When the cognitive effects (on mind and will to win) predominate over casualties inflicted because the accuracy and intensity of fire cause the enemy's accuracy and intensity to deteriorate, under those circumstances effective unit firepower is more valuable than numbers engaged. The losing side will discontinue the battle because it is being *dominated* by the enemy's fire, will over time be reduced to impotence, and will eventually be destroyed if it does not concede the enemy's military objective. Under such circumstances the traditional Lanchester square law conclusion favoring numbers over quality is reversed, and unit firepower is more influential than numbers of units.

A NUMERICAL EXAMPLE

As a numerical example, let A(0) = 10 units of force, and B(0) = 5 units. Let the initial destruction rate of A be α (0) = 1 kill of B per A unit per hour; and of B, β (0) = 2 kills of A per B unit per hour. The two sides are equal at the outset in the sense that

$$\alpha A = \beta B = 10 \textit{ kills per hour.}$$

If the Lanchester square law applies, unit fire is constant and continuous, with α = 1 and β = 2, and equation [2] obtains.

- A will win a fight to the finish, and when the B force is destroyed, A(T) = 7.07 units, so 71% of the A force survives. Such is the power of numbers under Lanchester square law conditions.
- If B senses defeat after one of its 5 units (20%) is lost and successfully ends the battle, A(T) = 9.05 when it is over. A has lost nearly one unit to B's superior fire, which is almost as much as B lost. A's long term cumulative advantage of numbers has not had time to take effect.

Next assume that the effect of suppression on enemy fire works much faster and is far more important than attrition. Equation [6] applies, so that in a formal sense neither side loses units. A = 10 and B = 5 are constants, and only the volume and accuracy of each side's fire is affected by the other's. Let the initial firing rate of A be a(0) = 10 rounds per minute per shoooter and let each round suppress enemy fire at s = .01 rounds per minute per enemy unit. Correspondingly, let b(0) = 20 rounds per minute with the same value of s. The two sides' initial volumes of fire are equal in that Aa(0) = Bb(0) = 100 rounds per minute.

- At time T when B has totally suppressed A's fire, each unit of B has an unsuppressed rate of fire remaining of b(T) = 1.414 rounds per minute, or 71% of the initial rate. We may presume A to be impotent and at the mercy of B.
- If A senses that B's volume of fire is dominant after his fire has been curtailed by 20%, he may end the battle and withdraw as he is able. At the time the battle is over, a(T) = 8 rounds per minute per shooter, and b(T) = 18.1 at its end, which is 90.5% of B's initial firepower.

INTERPRETATION OF THE SUPPRESSION MODEL

An interpretation of the suppression computation is interesting in several regards.

First, by supposition B inflicts no casualties, and so all of A may withdraw and live to fight another day. It is possible that A's force is permanently disorganized and demoralized, but that is not inherent in the model.

Second, the battle is won by superior unit firepower, even though at its outset the total rate of fire and fire effectiveness of both sides were equal; unit effectiveness is more influential than the number of units.

Third, B wins by forcing A to concede B's tactical objective or else suffer complete suppression and, by presumption, destruction. Mission accomplishment is a good way to decide who won the battle, but it is not the usual way in analysis, which is to compare casualties. Probably B's organization and morale will be stronger after the battle than before it, but again, the model does not tell us.

Fourth, we should not conclude that A's situation is hopeless in future battles. Now that we know the significance of treating a and b (or α and β) as time-dependent variables, we may anticipate battlefield conditions in which A finds a stronger position with an improved firepower ratio and so is able to exploit its numerical advantage.

Fifth, observe that the victory went to the side whose fire *dominated* on the battlefield. In a formal sense this was true by postulation, for we assumed that attrition played no part. Nevertheless it is useful to look at how and why the lethal potential of B's superior fire was decisive. It was in part because B's fire attenuated A's fire more rapidly, and in part because after A is reduced to impotence he must surrender or face destruction. In war the activation and effective employment of superior firepower is the central cause of victory, whether or not casualties determine the outcome. (Our model ignores movement for the sake of analytical simplicity, but it is safe to say battlefield movement unsupported by covering, suppressive fire is a rare occurrence. Tactical maneuvering is achieved by an astute blend of covering fire and movement.)

Sixth, these models are formalisms, whether basic equations [1], [3] or [5] are used. For one thing, casualties will occur on both sides. For another, the reduction of fire on the losing side caused by the winner's fire will usually not go all the way to zero. There is a point of diminishing returns in the suppression effect of fire towards attenuating enemy fire.

Seventh, as a reminder, ground combat is our subject. Naval and air combat take their own form. In the judgment of the author, attrition is the essential phenomenon, and suppression as a driving cause of sea and air battle outcomes is rare.

SUMMARY AND CONCLUSIONS

Let us summarize the major points of the quantitative analysis:

- The coefficients of unit effectiveness ought to be regarded as variables in time.
- Fire volume and unit effectiveness will often be far more variable than the change due to casualties, in which case they have a greater immediate effect on the outcome than attrition.
- When unit firepower is diminished more rapidly than the surviving number of units (due to the effectiveness of enemy fire) then the unit firepower advantage has the square law effect, not numerical advantage.

It is quite reasonable to expect that the suppression effect of superior combat power will frequently be greater than the attrition effect. Referring again to McQuie, at the end of 80 battles in and after World War II, the median casualties were 4% at the time the attacker abandoned his objective and 8% when the defender conceded defeat. McQuie cites the decisive effect of maneuver (or by implication the inability to maneuver) as the salient reason for termination by the loser in the majority of the battles. We have not examined the suppression of maneuver, though it is possible to do so. In any case, McQuie's data suggest that casualties are seldom the determinant of battle outcome. They also suggest that when the attacker's success results in a rout or surrender of the enemy, casualties may be principally in the form of prisoners taken instead of bodies slain.

One should not go so far as to disregard attrition. But even if one regards casualties as the available and hence utilitarian measure, nevertheless the Lanchester square law does not follow with its conclusion that numerical advantage is more valuable than unit firepower. Indeed, in recent research to demonstrate the utility of his "defender's advantage parameter," R. Helmbold [1995] shows that numerical advantage, which is the most important determinant in the square law, is among the worst predictors of victory in 83 historical battles. One should regard attrition and suppression both as important. This analysis suggests, to say the least, that their source—firepower—is more influential that numbers engaged.

But the answer we offer here to the question posed in the beginning is that sharply superior combat power must impose and exploit suppression when the goal is a relatively bloodless victory. The best example in the past 50 years of victory by suppression is the German blitzkrieg of 1940 in France. The blitzkrieg had many of the properties of the Gulf War. No attrition model can explain the blitzkrieg phenomenon, which was achieved by an intensive, local suppression of the defenders' ability to resist during a German armored-mechanized breakthrough. The breakthrough was followed by extensive demoralization of the French defenders when they faced an enemy in their rear. Blitzkrieg (a.k.a. lightning war or maneuver warfare) worked on the minds and spirits of the enemy to reduce his combat power to near zero.

It is in the nature of a successful campaign of maneuver that few battles are fought and the casualties that attract the interest of military historians and analysts are largely absent. To understand the full value of very superior combat power one must look for Sherlock Holmes' dog that didn't bark: mission attainment with the near-absence of combat and casualties. Perusal of battles analyzed quantitatively reveal few if any for the entire 1940 German campaign in France. For example, *Combat History Analysis Study Effort: CHASE*, by R. Helmbold [1986] lists only one, Sedan-Meuse River of 13 May 1940. Another example is the Japanese defeat of the British on the Malay Peninsula in 1941 in a brilliant campaign of maneuver with never a battle worthy of the history books. These are examples

of what Sun Tzu told us long ago is the true measure of successful generalship [1963, p. 77]. We recall what Clausewitz said in the same vein: it is not the attacker but the defender who initiates a battle; the defender may choose to concede the opponent his aim instead of fighting [1976, p. 377].

One may well ask why American military tacticians should not long ago have perceived the quite reasonable, if not obvious, relationships described above. It is probably a proclivity to leap at once to the "practical:" an ad hoc perspective and disinterest in general theory. Since 1950 the practical problem under study has been "how to fight outnumbered and win," tied to the NATO central front in Europe. The advantage of superior firepower, its potential for reducing casualties, and the deeper effects of the phenomena of combat power were not ours to exploit until the Gulf War.

The best military leaders have always sought victory by domination and control, not bloodshed. We analysts do a disservice if our studies cannot carry the quantitative analysis of battle outcomes beyond the measurement of casualties. The U. S. government seems to be in the process of casting aside the means to apply overwhelming combat power in the next major regional conflict. It would be most timely to give some attention to suppression as a root cause and dominance as the vital effect in determining the highly favorable outcome of the ground campaign in the Gulf War, and many others.

FUTURE RESEARCH

The Army and Marine Corps espouse "maneuver warfare." The thoughtful reader will see the possibility of extending the model herein in a number of ways that lead to a better understanding of maneuver's advantage, and the relationship between fire and movement.

a. The first and simplest is to work out the mathematics when both attrition and suppression result from fire; in other words, when fire reduces enemy numbers and return fire simultaneously. Some parametric analysis of the relationships will then be possible, but eventually the more difficult task of obtaining numerical values will have to be confronted. Clues to maneuver warfare are involved because one effect of suppressive fire is to "pin down" the enemy with one element of force (notably artillery and close air support), while another achieves a penetration (notably armor), or maneuvers on a flank (notably mechanized or other mobile forces).

b. The Lanchester square law form has been the basis of this discourse. There is no reason to think the square law is "right," except that aimed fire is the image of the hypothetical engagements described. I have been somewhat vague about the extent to which suppressive fire is aimed or fire-hosed. What we do know from Osipov, Helmbold, Hartley and others is that the square law advantage of numbers has rarely been achieved when measured by attrition. I think operations analysts of ground combat should wish to investigate whether historical results comport with combat models better after suppression is introduced. I do not know whether to suggest starting with the square law, the linear law, or something in between. Their instincts will be better than mine.

c. For the purpose of modeling maneuver warfare, a more direct approach is to extend the equations to include the suppression of movement as well as fire. On one hand, this is appealing because we know that fire inhibits enemy movement, yet combat models do not reflect it as a function of the accuracy and volume of enemy fire. On the other hand, one discovers a dimensionality problem and must wrestle with inputs in terms of the deceleration rate, which has units of meters/minute-enemy shot or the equivalent. If we analysts are serious about helping to explore the operations and tactics of maneuver warfare, it seems incumbent on us to try, by showing the connection between movement and supporting (suppressing) fire in maneuver warfare. For what it is worth, it is certainly possible to write equations in which fire has all three effects: casualties, suppression of fire, and suppression of movement.

d. Another approach would be to regard the effect of opening fire as a step function. In such a model, the instant fire is opened, enemy return fire (or motion) is diminished to a new, lower level. This is more or less implicit now in the inputs for the coefficients of unit effectiveness. The new feature would be to scale the

effectiveness of B's fire (or motion) inversely with the accuracy and intensity of A's fire. To my knowledge, this has never been attempted. The conceptualization is not a big step, but the scaling factors would come only with great difficulty.

ENDNOTES

1. To TMCI it seemed desirable to call their cornerstones Axioms, whether traced from an innate comprehension of truth or derived from common experience. The six axioms are thought to be the fewest possible, so obvious as to be indisputable, and so essential that no structure of theory or model should proceed without cognizance of them. Root definitions must, of course, accompany them:

Definition 1. A MILITARY FORCE is a set of elements which are activated for combat.

AXIOM 1. COMBAT occurs by deadly interactions between military forces.

Definition 2. COMMAND is the function which organizes, motivates, makes decisions regarding, and directs the activities of its force.

AXIOM 2. In combat each COMMAND seeks to achieve a goal, called its mission, which has perceived value.

Definition 3. COMBAT POTENTIAL is the latent capacity of a military force to achieve results in combat.

AXIOM 3. COMBAT POTENTIAL is embodied in military forces.

Definition 4. COMBAT POWER is the realized capability of a military force at an instant of time to achieve results in combat.

AXIOM 4. Command activates its combat potential to create COMBAT POWER in furtherance of a mission.

Definition 5. DOMINATION is the condition wherein one military force imposes its will on the other.

AXIOM 5. In combat, DOMINATION of the opposing military force is the ultimate means of achieving an objective.

AXIOM 6. UNCERTAINTY is inherent in combat.

The status of TMCI theory of combat may be had from the author at the Naval Postgraduate School, Monterey, California 93943, or Dr. Donald S. Marshall, Executive Director of The Military Conflict Institute, 12 Fairfield Street, Salem, Massachusetts 01970.

2. A source of some rare exceptions is MORS' two-volume Proceedings of the Second Mini-Symposium on the topic, *Human Behavior and Performance as Essential Ingredients in Realistic Modeling of Combat—MORIMOC II.* Among the noteworthy papers are:

a. David Rowland, "Assessment of Combat Performance with Small Arms"

b. Charles L. Frame, Brian R. McEnany and Kurt A. Kadivko, "Combat Operational Data Analysis: An Examination World War II Suppression Data"

c. George Schecter, James C. Richards and Henry A. Romberg, "Tactical Deterrent Effects Model"

d. Trevor N. Dupuy, "The Fundamental Information Base For Modeling Human Behavior In Combat"

3. On page 2 of this early paper by one of our most prolific and perceptive analysts of historical ground combat, Helmbold writes: "[Daniel] Willard's approach is typical of those which attempt to infer the form of Lanchester's equations from an analysis of a large number of battles for which the initial and final strengths on both sides are known, without depending on any information about the details of their attrition histories. A finding that this approach is fully justified would be of capital importance for the development of a theory of combat, because data on initial and final strengths are available for hundreds of battles. . .they at least are much

more readily available than data on attrition histories" [1979]. In the preceding paragraph Helmbold has just finished taking note of the famous effort by J. Engel to validate the Lanchester square law's *shape* as it unfolded during the extended Battle of Iwo Jima, so he is clearly concerned with the attrition *history* occuring within a battle, or short campaign.

4. A more thorough description would backtrack to equations [3] to explain that functions α and β are really composites of a firing rate term in shots fired/minute, and an accuracy term, in hits per shot. Then in equations [4] we would have same firing rate value as in equations [3] plus a suppression term, in shots surpressed per shot fired.

5. In memorable and decisive battles, casualties have been as much as 25-35%, but such battles are usually episodic, comprising separate phases in which large changes of α and β in effect *define* the phases.

REFERENCES

J. V. Chase, "A Mathematical Investigation of the Effect of Superiority of Force in Combats Upon the Sea," Newport RI, Naval War College archives, 1902; reprinted in its entirety as Appendix C of B. A. Fiske [1988].

Carl von Clausewitz, *On War*, edited and translated by Michael Howard and Peter Paret, Princeton, NJ, Princeton University Press, 1976.

P. K. Davis and D. Blumenthal, "The Base-of-Sand Problem: A White Paper on the State of Military Combat Modeling," Santa Monica, RAND, N-3148-OSD/DARPA, 1991.

T. N. Dupuy, *Understanding War: History and Theory of Combat*, New York, Paragon House, 1987.

B. A. Fiske, "American Naval Policy," *Naval Institute Proceedings*, April 1905.

D. S. Hartley III, *Confirming the Lanchestrian Linear-Logarithmic Model of Attrition*, Report K/DSRD-263, Oak Ridge, TN, Martin Maritta Energy Systems, Inc. December 1990.

D. S. Hartley III, "Validating Lanchester's Square Law and Other Attrition Models," *Naval Research Logistics* Journal, June 1995.

D. S. Hartley III, *Predicting Combat Effectiveness*, Report K/DSRD-412, Oak Ridge TN, Martin Marietta Energy Systems, Inc. August 1991.

R. L. Helmbold, "Some Observations on the Choice of Exponent in Lanchester's Law," copyrighted manuscript, June 1979.

R. L. Helmbold, *Combat History Analysis Study Effort (CHASE): Progress Report For the Period August 1984-June 1985*, Bethesda, Md. U. S. Army Concepts Analysis Agency, CAA-TP-86-2, August 1986.

R. L. Helmbold, "Combat Analysis: Defender's Advantage Parameter, Some Empirical Evidence," *PHALANX*, March 1995.

G. W. S. Kuhn, *Ground Forces Casualty Rate Patterns: The Empirical Evidence*, Report FP703TR1, Bethesda, MD, Logistics Management Institute, September, 1989, and *Ground Forces Battle Casualty Rate Patterns: Current Rate Projections Compared to the Empirical Evidence*, Report FP703TR2, May, 1990.

F. W. Lanchester, "Aircraft in Warfare: The Dawn of the Fourth Arm, Number V, The Principle of Concentration," *Engineering 98*, 1914 (Reprinted pp. 2138-2148 of *The World of Mathematics*, J. Newman, editor, New York, Simon and Shuster, 1956).

R. McQuie, "Battle Outcomes: Casualty Rates as a Measure of Defeat," *Army Magazine*, November 1987.

Military Operations Research Society, Proceedings of the second mini-symposium on *Human Behavior and Performance as Essential Ingredients in Realistic Modeling of Combat - MORIMOC II*, two volumes, Alexandria VA, MORS, May 1989.

M. Osipov, "The Influence of the Numerical Strength of Engaged Forces on Their Casualties," originally in the Russian Journal, *Military Collection*, June-October 1915; Translated and annotated by R. L. Helmbold and A. S. Rehm in research paper CAA-RP-91-2, Bethesda, Md, U. S. Army Concepts Analysis Agency, 1991.

Sun Tzu, *The Art of War*, Translated with an introduction by S. B. Griffith, New York, Oxford University Press, 1963.

The George Washington University
WASHINGTON DC

EIGHTEENTH SYMPOSIUM ON MATHEMATICAL PROGRAMMING WITH DATA PERTURBATIONS
23-24 MAY 1996

CALL FOR PAPERS

The EIGHTEENTH Symposium on Mathematical Programming with Data Perturbations will be held at George Washington University's Marvin Center on 23-24 May 1996. This symposium is designed to bring together practitioners who use mathematical programming optimization models and deal with questions of sensitivity analysis, with researchers who are developing techniques applicable to these problems.

CONTRIBUTED papers in mathematical programming are solicited in the following areas:

(1) Sensitivity and stability analysis results and their applications.
(2) Solution methods for problems involving implicitly defined problem functions.
(3) Solution methods for problems involving deterministic or stochastic parameter changes.
(4) Solution approximation techniques and error analysis.

"CLINICAL" presentations that describe problems in sensitivity or stability analysis encountered in applications are also invited.

ABSTRACTS of papers intended for presentation at the Symposium should be sent in triplicate to Professor Anthony V. Fiacco. Abstracts should provide a good technical summary of key results, avoid the use of mathematical symbols and references, not exceed 500 words, and include a title and the name and full mailing address of each author. The deadline for submission of abstracts is 17 March 1996.

APPROXIMATELY 30 minutes will be allocated for the presentation of each paper. A blackboard and overhead projector will be available.

ANTHONY V. FIACCO, ORGANIZER
Sponsored by the Department of Operations Research and the Institute for Management Science and Engineering, School of Engineering and Applied Science
The George Washington University, Washington. D.C. 20052
Telephone: (202) 994-7511

RIST PRIZE CALL FOR PAPERS

MORS offers two prizes for best papers—the ***Barchi Prize*** and the ***Rist Prize***. The ***Rist Prize*** will be awarded to the best paper in military operations research submitted in response to this **Call for Papers**. The ***Barchi Prize*** will be awarded to the best paper from the entire 64th symposium, including Working Groups, Composite Groups, and General Sessions.

David Rist Prize: Papers submitted in response to this call will be eligible for consideration for the **Rist Prize**. The committee will select the prize-winning paper from those submitted and award the prize at the 65th MORSS. If selected, the author(s) will be invited to present the paper at the 65th MORSS and to prepare it for publication in the MORS journal, *Military Operations Research*. The cash prize is $1000. To be considered, the paper must be mailed to the MORS office and postmarked no later than **September 29, 1996.** Please send the original, three copies and the disk.

Richard H. Barchi Prize: Author(s) of those papers selected as the best from their respective Working Group or Composite Group, and those of the General Sessions at the 64th MORSS will be invited to submit the paper for consideration for the ***Barchi Prize***. The committee will select the prize-winning paper from among those presented, nominated and submitted. The prize will be presented at the 65th MORSS. The cash prize is $1000. To be considered, the paper must be mailed to the MORS office and postmarked no later than **November 30, 1996.** Please send the original, three copies and a disk.

PRIZE CRITERIA

The criteria for selection for both prizes are valuable guidelines for presentation and/or submission of any MORS paper. To be eligible for either award, a paper must, at a minimum:

- Be original and a self-contained contribution to systems analysis or operations research;
- Demonstrate an application of analysis or methodology, either actual or prospective;
- Prove recognizable new insight into the problem or its solution; and
- Not previously been awarded either the Rist Prize or the Barchi Prize (the same paper may compete for but cannot win both prizes.)

Eligible papers are judged according to the following criteria:

Professional Quality

- Problem definition
- Citation of related work
- Description of approach
- Statement of assumptions
- Explanation of methodology
- Analysis of data and sources
- Sensitivity of analyses (where appropriate)
- Logical development of analysis and conclusions
- Summary of presentation and results

Contribution to Military Operations Research

- Importance of problem
- Contribution to insight or solution of the problem
- Power of generality of the result
- Originality and innovation

Managing the Pilot Force in an Uncertain Environment

Harry J. Thie
William W. Taylor
Claire Mitchell Levy
Clifford M. Graf II
Sheila Nataraj Kirby

RAND
2100 M Street NW, Suite 800
Washington, DC 20037-1270

INTRODUCTION

The services are undergoing a fundamental reshaping and restructuring, driven by the demands of a new era of tighter fiscal constraints, new security challenges, new technology, and increased reliance, in some services, on the reserve components. For example, in the last seven years, Air Force active end strength has decreased by over 180,000, and more reductions are proposed. By the end of FY 95, Air Force active end strength will total about 400,000, while reserve strength will remain approximately the same at about 194,000.

To reduce strength, the services are constraining accessions, encouraging voluntary departures, and imposing involuntary separations. This paper focuses on one personnel group—pilots in the Air Force and the Navy—and examines how service personnel policies that affect pilots are changing in the wake of downsizing and restructuring. In particular, we provide a critical assessment of three areas: (1) the requirements for pilots in the context of the defense drawdown and restructuring; (2) the supply of pilots in the same context and the sustainability of such a force—given historical trends in accession, retention, and transfers to the reserves; and (3) the effectiveness of current personnel management and training policies in terms of their effectiveness in meeting future needs.

HISTORICAL OVERVIEW

In order to set the current situtation in context, it might be helpful to review some historical trends in pilot requirements, supply, and management using the example of the Air Force. This would provide some context against which the uncertainties and changes that are occurring now can be evaluated.

As Figure 1 makes clear, there has been constant change over the past forty years, both up and down. In general, the pattern that emerges is that requirements appear to change more quickly at the beginning of a period of growth or decline, but inventory appears to lag behind. The services cannot develop or separate military pilots as quickly as their needed numbers can be changed on paper. During military buildups, the inventory of pilots typically falls short of requirements. This was the case during the Korean War of the 1950s, the Vietnam Conflict in the 1960s, and the Reagan build-up of the 1980s. As forces are reduced during post-conflict eras, inventory tends to exceed requirements.

Recognizing that supply will always lag requirements, the policy question is how to bring supply and demand closer to equilibrium. There have been two traditional long term policy "balancers." First, increases or decreases in accessions (training of new pilots) can be used to bring supply (inventories) into equilibrium with demand

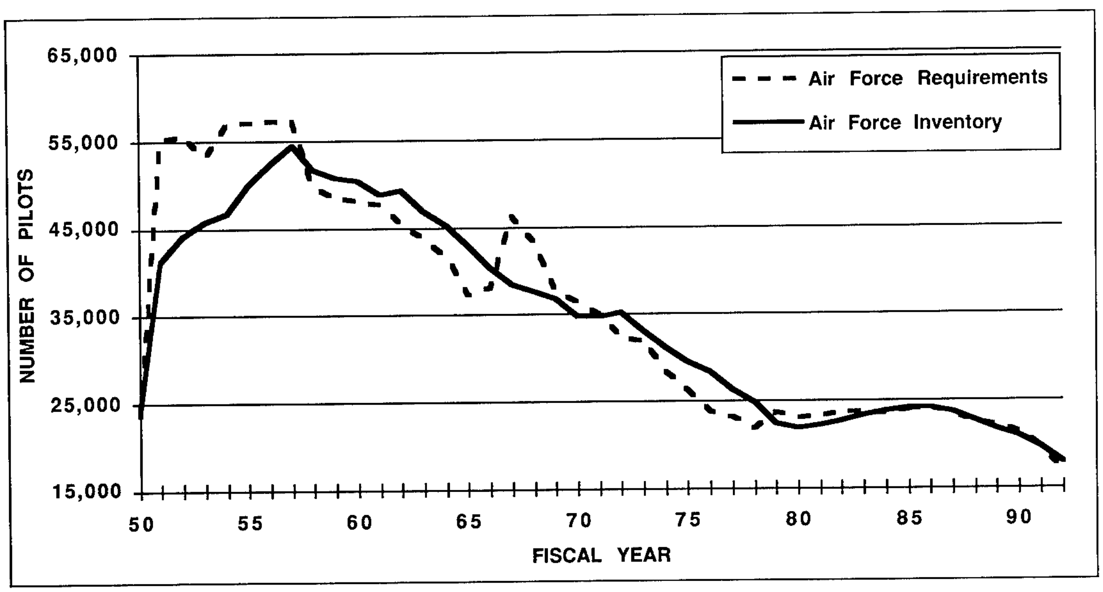


Figure 1: Historical Pilot Management--Air Force Pilots 1950-1992

(requirements). However, when one examines the Air Force's pilot requirements with Air Force Undergraduate Pilot Training (UPT) accessions over a forty year period (Figure 2), it is clear that changes in demand (requirements) are not immediately met by changes in training. In particular, reductions in demand are difficult to match because of the training pipeline. Commitments to UPT are typically made by undergraduate students several years before they even begin training. As can be seen, there was a dramatic reduction in the requirements for pilots in 1967, as the Vietnam War began to draw down. However, reductions in UPT accessions did not begin until 1972.

A second long term policy "balancer" has been retention. The measure of effectiveness used for retention in this chart is the Air Force's Cumulative Continuation Rate (CCR),[1] a measure of retention in the 6 to 11 year group (the critical time after the Active Duty Service Obligation ends and before officers make career decisions to stay through to retirement). This measure has only existed since 1976, the time when the Air Force began its Rated Management process.

In the short-run, the Air Force has traditionally used the assignment system to mitigate shortages and surpluses. A priority assignment system, where certain requirements are filled on a priority basis with some other requirements left unfilled, has been used to cope with historical shortages of pilots. During the surplus period in the 1980s, the "rated supplement" allowed the assignment of rated officers not needed for pilot assignments to non-rated career fields for career development. This also provided a surge capability to meet contingency requirements in the past. (This contingency role is now filled by reserve components.)

The Navy uses the active-duty reserve to help balance supply and demand. Increasing pilot requirements are met by increasing the number of newly commissioned U.S. Naval Reserve (USNR) officers to enter flight training. There was a marked decline in the number and proportion of active duty reserve pilots from the late 1960s-early 1970s to the early 1980s. In 1971-72, approximately 35 percent of pilots were active duty reserves. However, as requirements declined, these pilots were released from active duty or not brought on active duty as was the

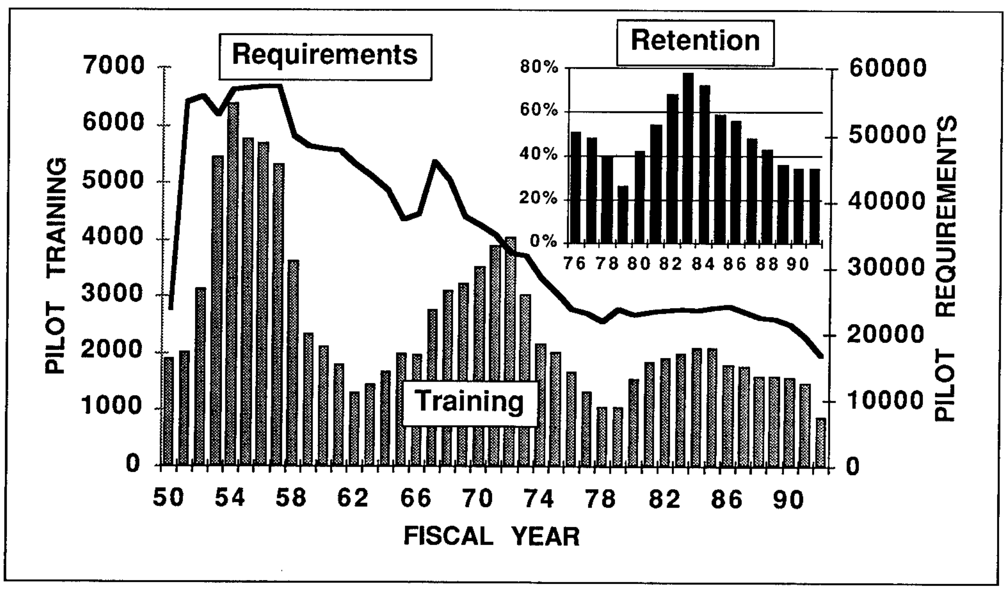


Figure 2: Traditional Long Term Policy "Balancers" Are Training and Retention

case between 1972 and 1980. By 1980, only 13 percent of pilots were active duty reserves. This proportion rose from 1980 to 1988, when active duty reserves comprised about 30 percent of pilots. Currently, 23 percent of pilots are active duty reserves. The Navy is separating USNR pilots, who would be willing to stay, from active duty to accommodate the ongoing drawdown.[2]

CURRENT CONTEXT

We are currently in the midst of a period of dramatic change, largely driven by the post-Cold War drawdown.

First, forces are being reduced significantly. The first iteration of reductions were to a Base Force level by 1995; the next iteration is to a Bottom Up Review level by 1999. With respect to aviation, the effects on the Air Force thus far have been the greatest. By 1992, they had undergone the bulk of their cuts in building down to Base Force Levels, having already absorbed approximately two thirds of these cuts. In general, the Navy drawdown began later than the Air Force drawdown. But, as with the other services, the post-Cold War era will have a significant impact on the Navy. The new strategic direction will affect naval aviation in three basic areas—force structure, organizational structure, and policy about integration of Naval and Marine Air. Even now, the Navy continues to assess the overall number of squadrons needed for the defined force.

Second, the Air Force, (and to a lesser extent the Navy), is undergoing a major reorganization. The Air Force reorganization which is complete on paper and which they began implementing in 1992 affected all levels from Air Staff to the Wing. The Air Force continues to assess the need for certain staff positions in its component organizations. In the Navy, squadrons and air wings will differ from their predecessors in that the mix of aircraft will be different (e.g., the A-6 will leave the inventory), the number of aircraft will be different (i.e., 50 plane wings instead of wings of 80 planes or more), and Marine Air will be more closely integrated (e.g., active and reserve Marine squadrons operating as part of CVWs).

Third, there are force structure shifts between active and reserve components, that will have important impacts on requirements. The Naval Air Reserve will reduce to one carrier air wing from two. The Air Force active-reserve force mix is changing. The active Air Force is being reduced significantly; the Air Reserve Component (ARC) less so. This will result in a growing share and changing mix of aircraft and missions being housed in the ARC. Additionally, policy changes in the active component have a future effect on the ARC. For example, the decrease in active flying units and the excess supply of separating active pilots in the near-term allows good recruiting of prior service pilots by the ARC now, but the opposite will occur in the future. Extending the active-duty service commitments after completion of undergraduate pilot training will change the future dynamics of reserve pilot flow. Retention bonuses for pilots who agree to serve through their 14th year of service also affect the number and timing of active separations. Increases in active service duty will age the future pool of prior service pilots available to the ARC, something that the ARC may not view as a desirable shift.

It is clear that new ways of thinking and new policies will be needed to deal with this changed and still changing environment.

PILOT REQUIREMENTS

We developed three classes of determinants—force structure, organizational structure, and other policy change (A more complete analysis is available from the authors). Force structure changes center on the drawdown to include fewer primary aircraft authorized (PAA). Organizational restructuring covers the full spectrum of internal reorganization that the services are undergoing from squadron and wing through service headquarters. In addition, there are other policy changes which impact on pilot requirements such as changes in crew ratios.

We adapted Air Force terminology and divided pilot requirements into four basic categories—force, staff, training, and manyear or other which includes transients and students. Force requirements include all of the pilots required to man operational flying units. Training programs are necessary to ensure the operational readiness and necessary skill level of the operational units. This category includes the instructors for both formal and other training. Staff

are necessary in order to provide for overhead support and supervision required to ensure safe, efficient, and productive flying operations. Finally, there are other or manyear requirements which allow for time in the training pipeline, enable pilots to participate in professional development or formal education programs, or permit pilots to take leave in conjunction with reassignment.

Since we ultimately want to compare future inventories of pilots with requirements for them for the Air Force and Navy, we used our evaluation of Air Force and Navy pilot requirements to estimate objective profiles for those requirements in the steady state (FY 97+). These profiles convert data by grade to data by years or length of service. We use them as an objective that policies for pilot management and training are trying to achieve. Such policies cause changes in inventory profiles which can then be compared to the requirements profiles to see whether and if the objectives are being met. The FY 97 requirements projection was made taking into account the Bottom-Up Review force levels.

PILOT INVENTORIES

Pilot inventories are determined by production and retention. Inventories are also influenced by other factors to include: distribution, absorption capacity, assignment and promotion policies. We provide brief definitions of each of these terms.

Production. The annual rate of graduates from Undergraduate Pilot Training (UPT). This factor is closely related to UPT *accessions*; the differences are due to attrition during UPT.

Retention. A measure of the likelihood of pilots to continue from year to year in the pilot inventory. This factor is determined by *loss rates*, which give the percentage of a cohort population that is lost from the inventory in a given year. Forced retention has traditionally been legislated in the form of a minimum active duty service obligation (ADSO) following the completion of flying training. The current ADSO is eight years. Voluntary retention can be influenced by promotion, compensation (such as Aviation Career Incentive Pay and Aviation Continuation Pay), quality of life, and other factors such as airline demand (together with salary and job security forecasts).

Retention is one of the key factors affecting pilot inventory. One of the central problems for retention of military pilots is civilian airline hiring. Our analysis (Mitchell, 1995) shows that in the near term, there are plenty of military (and other) pilots to satisfy civilian airline demand through the end of the drawdown in 1997, driven primarily by two factors. On the demand side, the airline industry will continue to hire at levels well below historic averages through the end of this decade. On the supply side, pilots will continue to separate from the military through the drawdown in order to reduce our forces to their post-Cold War levels. After the year 2002, as airline demand increases, and the reserve pool for hires becomes much smaller, we believe a situation of excess civilian demand will exist and may well cause problems for military pilot retention.

Distribution. The assignment of new (and other non-MWS identified) pilots to actual pipeline training which establishes major weapons system (MWS) credentials.

Absorption capacity. The number of new and previously qualified pilots that each MWS group can accept each year. This factor is influenced (or constrained) by several policy issues and other parameters. These include:

(a) *Experience:* two components are incorporated in this parameter. The first establishes criteria in terms of minimum flying time and/or time in crew position required for a crewmember to be identified as experienced in a given MWS. The second determines minimum proportions of aircrew authorizations (by flying unit) which must be filled by crewmembers who meet these criteria and are thus identified as experienced.

(b) *Stability goals:* This includes (i) Assignment stability: length of time at one base prior to a reassignment; (ii) Weapon system stability: length of time flying a particular weapon system; (iii) Aircraft Commander stability: length of time in aircraft after upgrade to aircraft commander; (iv) Primary aircraft authorizations (PAA) and crew ratios (by MWS): all absorbing cockpits are part of the Crew Ratio Force; (v) Vacancies in absorbing units, or the number of absorbing cockpits; (vi) Pipeline (i.e. post-UPT) training capacity (by MWS); (vii) Loss rates and retention (which are, in turn, impacted by a number of other factors).

Absorption capability clearly influences (and constrains) distribution and number of new pilots to be produced each year.

Assignment and Promotion Policies. These factors have a significant impact on absorption ability and retention. Officers can also be promoted out of the pilot inventory, which only includes grades O-1 through O-5.

Interaction of the Factors. These factors dynamically interact to complicate real world inventory management. For example, the Air Force has moved over time from accessing pilots to sustainment levels, to accessing pilots to absorption ability, and finally to a vacancy driven system as the factors have changed.[3]

COMPARING PROJECTED DEMAND WITH SUPPLY

Figure 3 compares our estimate of Air Force FY 97 inventory with requirements both in the aggregate and by MWS (See Thie et al, 1994 for a more complete discussion). There are no *critical* shortfalls anywhere prior to FY 97, although shortages do start to occur in FY 97. By critical, we mean less than 5 percent short and able to be resolved without major policy initiatives. There are some minor imbalances by MWS.[4]

A similar assessment for the Navy produces a similar finding (Figure 4). Indeed, the Navy situation in FY 97 appears to be significantly improved over that of FY 92 when significant numbers of helicopter pilots and Naval Flight Officers were required to fill the Aviation Generalist (15XX) billets. These projections are based on Navy generated UPT production rates (which include MWS distributions) through FY 97.

The next two figures (Figures 5 and 6) provide a year of service (YoS) comparison of the FY 97 inventory projection with our estimated FY 2002 requirements objective profile for each service. The comparisons confirm a significant experience maldistribution, the shortfalls occur in the initial cohorts in the profile (YoS 2 through YoS 5), while corresponding overages are generated in the YoS 7 through YoS 10 cohorts. We confirmed similar experience profiles for the individual MWSs. This experience maldistribution generates no immediate operational problem, since shortages occur among the least experienced (and presumably least capable) pilots, and the overages occur among those cohorts which typically represent the heart of operational flying experience (i.e. the vast majority of instructor pilots, flight examiners, and

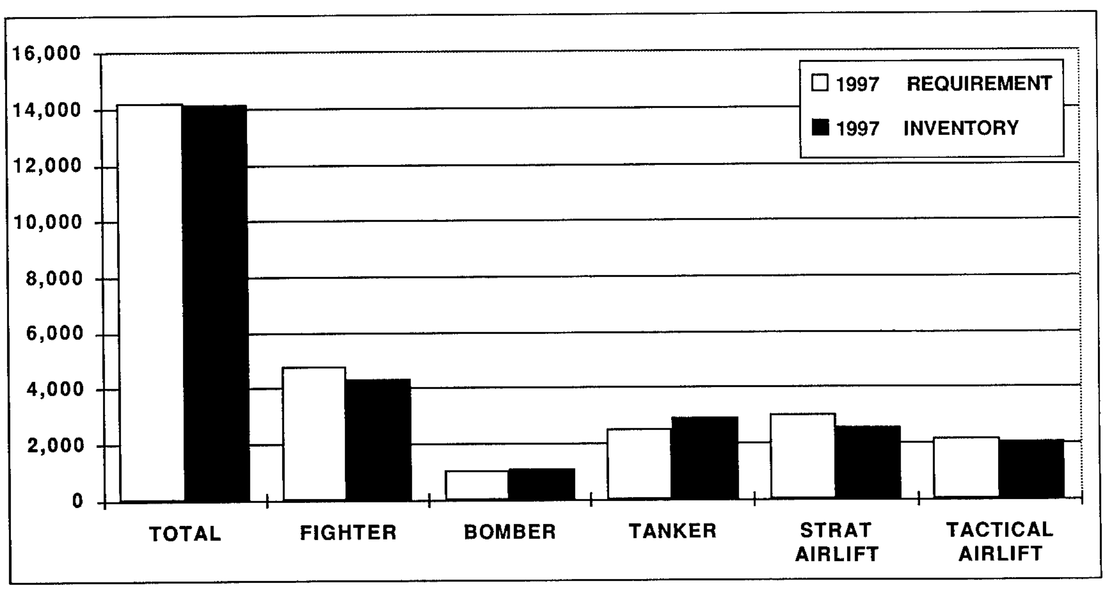


Figure 3: No Critical Numerical Shortages Exist in Any Air Force MWS Through FY 1997

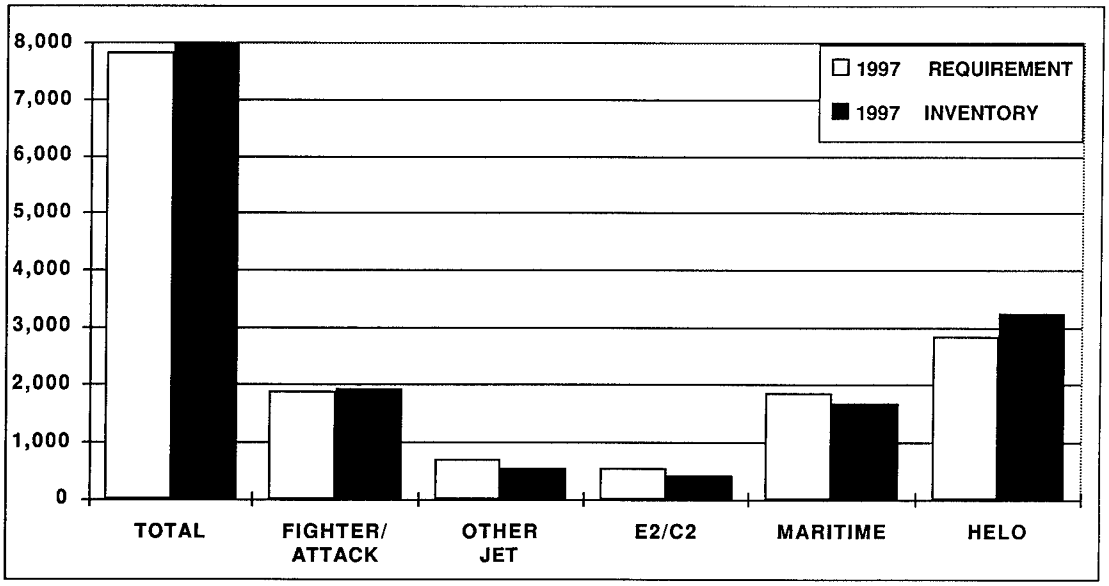


Figure 4: No Critical Numerical Shortages Exist in Any Navy MWS Through FY 1997

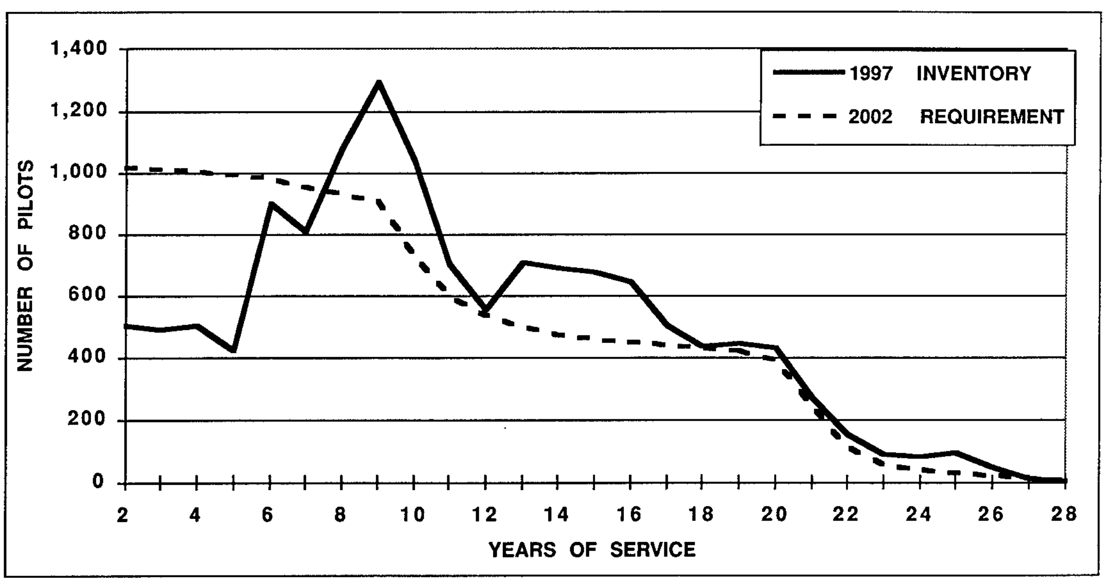


Figure 5: Air Force Experience Maldistribution in FY 1997 Is No Problem

other inflight supervisors are taken from these Air Force Captain/Navy Lieutenant cohorts).

The Navy's projected experience distribution for FY 97 conforms somewhat closely to the requirements objective profile than does that of the Air Force. Ironically, the Navy's higher attrition experience prior to ten years of service (due to shorter active duty service obligations and lower natural retention) allows the Navy to avoid the significant overages encountered by the Air Force in the YoS 7 through YoS 12 cohorts and enables the Navy to maintain reasonable UPT production levels throughout the drawdown period which will help avoid future cohort shortfalls.

A complete assessment, however, must also consider the dynamic behavior of the inventory. What is the potential operational impact as the reduced cohorts age to where they would normally assume increased flying responsibilities? To examine this issue, we continued our inventory projection model for another five years to the year 2002.

For the Air Force, Figure 7 confirms that the small accession cohorts of the early and mid-1990s will generate a "bathtub" effect in the inventory by FY 2002, and it also establishes that inventory shortfalls will occur by that time. The shortfall shown represents some 1,400 Air Force pilots. Individual MWSs reflect similar patterns to those exhibited by the total inventory, though these patterns are less pronounced in every MWS except fighters.

Alignment of inventory and requirement distributions for the Navy in FY 2002 is remarkably close (Figure 8). Indeed, an overall 200 pilot surplus is a significant improvement over the pre-drawdown situation in FY 92 in which a numerical shortage of 2,000 pilots existed. (Some drawdown losses probably occurred in FY 92 inventory that were not reflected in requirements.)

The conclusions are somewhat conditional, because as we point out above, the severity of problems in the outyears will be critically dependent on (1) requirements and (2) the potential policy options which might be employed to counter the indicated problem areas. We need to examine what options are available and how successful they might be in mitigating or alleviating problems. We first address the inventory shortfall problem. We focus on the Air Force for the remainder of our analysis. Many of the policy options, of course, apply equally to the Navy, although we do not forecast problems for them.

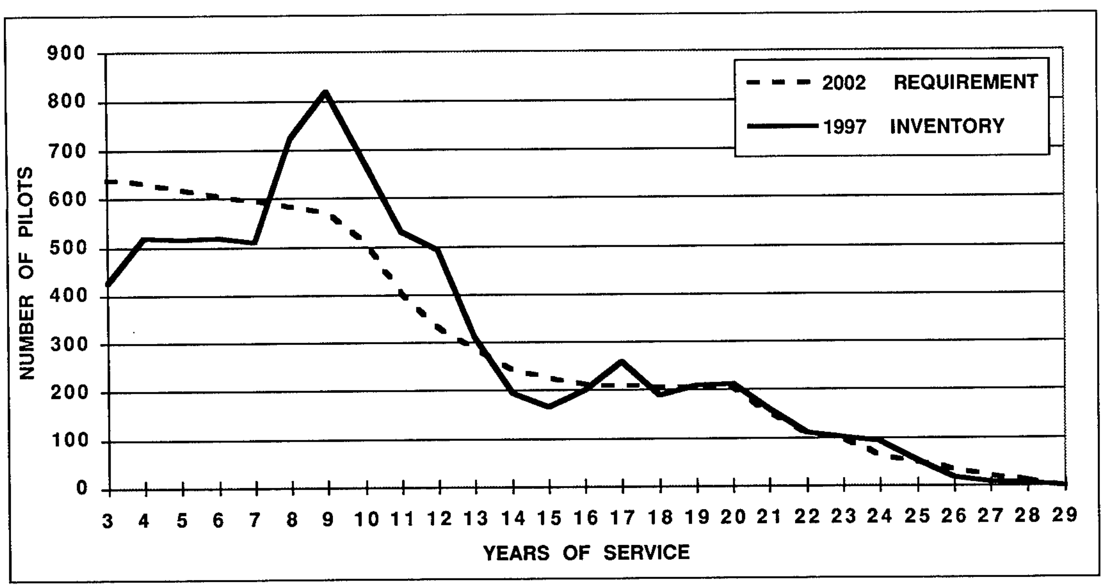


Figure 6: Navy FY 1997 Experience Distribution Shows No Major Problem

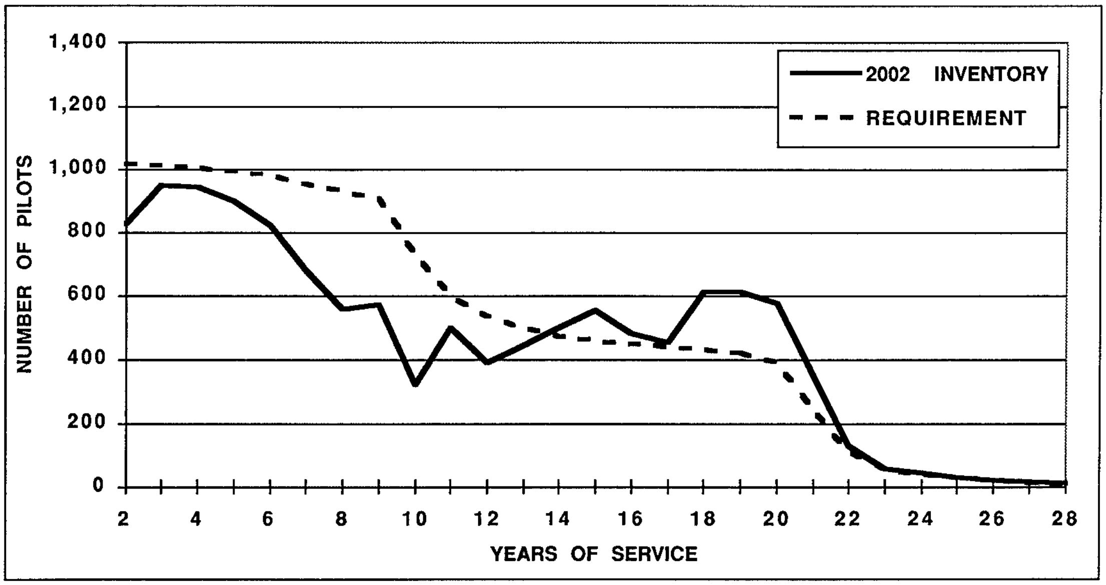


Figure 7: In 2002, Air Force Has a "Bathtub" and a 1,400 Pilot Inventory Shortfall

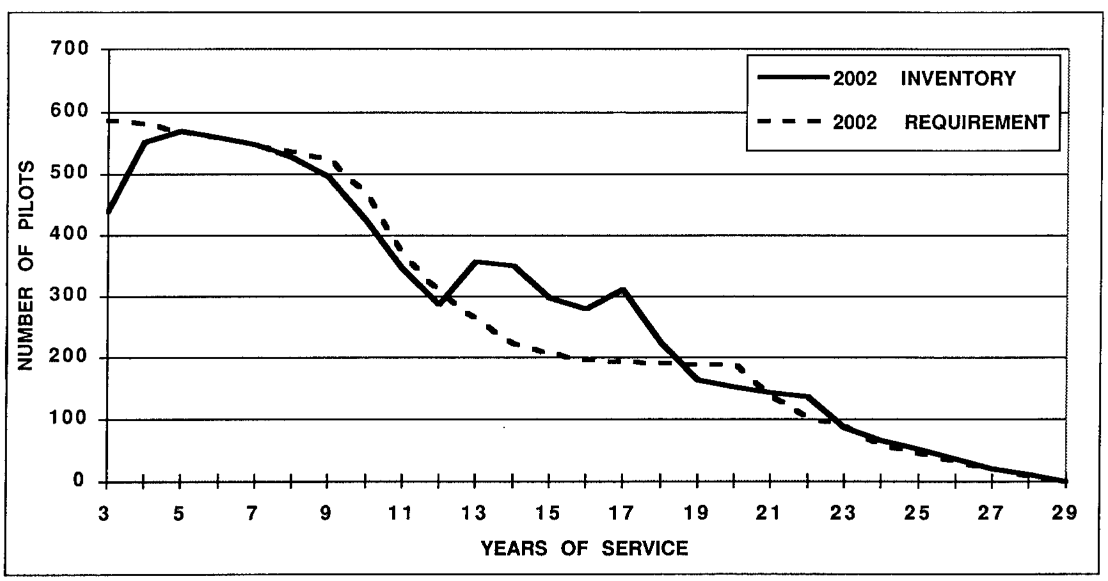


Figure 8: In 2002, Navy Has No Major Problems

PILOT MANAGEMENT AND TRAINING

We had mentioned earlier that there are two primary "balancers": training and retention. Clearly, if we are projecting a shortfall of pilots in the future, the two policy choices that we need to look at is training larger numbers of new pilots and retaining more experienced pilots. The discussion below outlines several alternative policy options and examines the impact of each on the problems identified in the previous section.

EFFECTS OF CHANGED RETENTION POLICIES

The first two alternatives we examine are based on changed retention. In order to build what could be considered a "control" or comparison group, we first characterized a base case. Retention is measured in terms of Years of Total Active Rated Service or TARS, which is the expected number of years of rated service that an average pilot will serve on active duty after completing UPT.[5] For the base case, we assume the following:

(1) FY95-FY97: high retention—13.5 TARS;
(2) FY98-FY99: lowest retention—12.0 TARS;
(3) FY00-FY01: low retention—12.5 TARS.

The rationale for these assumptions is discussed below.

Pilots reaching the retention point in FY95-FY97 are those that have been "shaped" by the drawdown and the officers remaining should be a select group. Thus it seems credible to assume that when they approach the retention window, they should have higher retention than might typically be the case. Figure 9 illustrates what we mean by a shaped cohort. The dotted line shows the FY93 year of service requirements profile matched against the FY93 beginning year inventory, shown by the full height of the bars. Because of the large overages in the more experienced year groups, the Air Force instituted a series of force reduction policies targeted at these groups. For example, the Pilot Early Release Program was targeted at those in the YoS 6-14 groups, Variable Separation Incentives/Selective Separation Bonuses (VSI/SSB) were offered primarily to those in YoS 14+ groups, while Selected Early Retirement (SERB) was mandated for some in the YoS 24+ groups. As a result of the reductions (shown by the cross-hatched, lighter, or darker areas at the top of the bars), the FY93 end year inventory more closely matched the required profile, although discrepancies are still evident. This provides the rationale for the

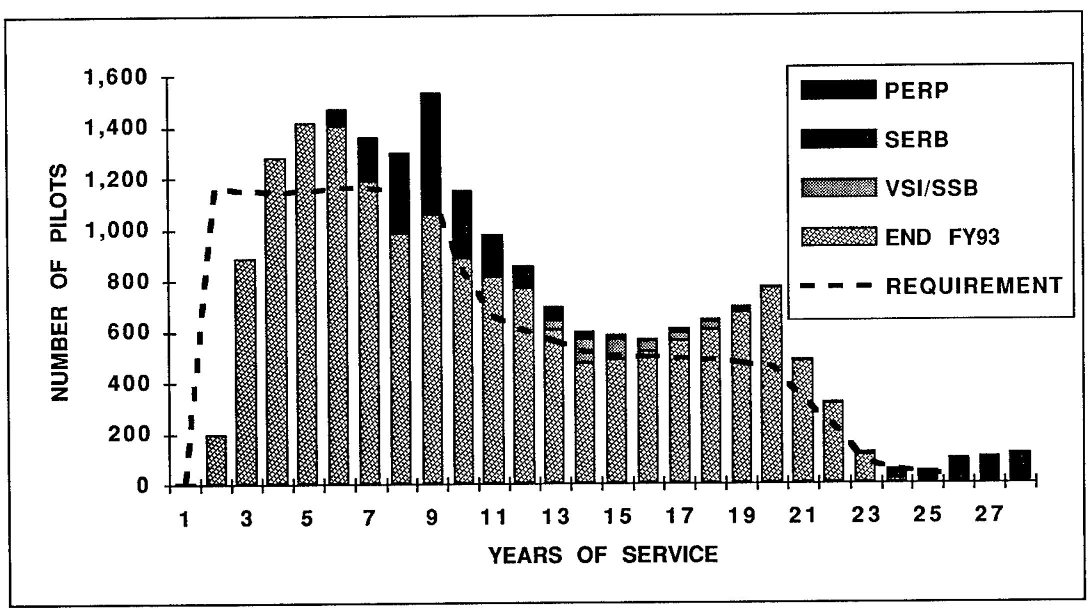


Figure 9: Example of a Shaped Cohort

assumption of high retention for these cohorts in the base case described above. Many of those who might have left later have already done so earlier.

The number of pilots reaching the retention point in FY98-FY99 cohorts is reasonably large while the number in FY00-FY01 is smaller because fewer pilots were trained in the early 1990s; the generally observed inverse relationship between cohort size and retention provides the rationale for the particular assumptions we adopt for these cohorts.

Table 1 outlines Alternatives 1 and 2 (as well as the base case for contrast) which bound the realm of possibilities with respect to changed retention. Alternative 1 assumes uniformly high retention for all cohorts, regardless of size or year (13.5 TARS); Alternative 2 assumes a much lower retention for the shaped cohorts (12.5 instead of 13.5 TARS), while maintaining the assumptions regarding retention of the FY98-FY01 cohorts (12.0 and 12.5 TARS respectively).

The results of our analysis under the different alternatives are displayed in Figure 10. The year of service requirement profile is shown by the dotted line, while the black bars depict the FY02 inventory under the base case. As we had shown earlier, aggregate requirements in FY02 are projected to be 13,700 and we estimate a shortfall in the base case of 1,400 pilots. The bathtub is very evident. What is interesting and perhaps surprising is that changed retention appears to only marginally affect the inventory profile. In the aggregate, we estimate that increased retention (Alternative 1) will increase the overall inventory to 12,900, thus reducing the estimated shortfall by 600 pilots while lower retention (Alternative 2) will decrease the overall inventory to 11,700, thereby increasing the shortfall by 400 pilots. Note that there is, however, no measurable effect on the maldistribution of experience. The different retention scenarios affect primarily those in the more experienced year groups, YoS 10-18, with little or no effect on the bathtub.

EFFECTS OF CHANGED TRAINING POLICIES

The next alternatives focus on increased training. The base case is drawn from Air Force plans that were in effect at the time of the analysis. Subsequent to that, the Air Force, recognizing the future value of increased UPT now,

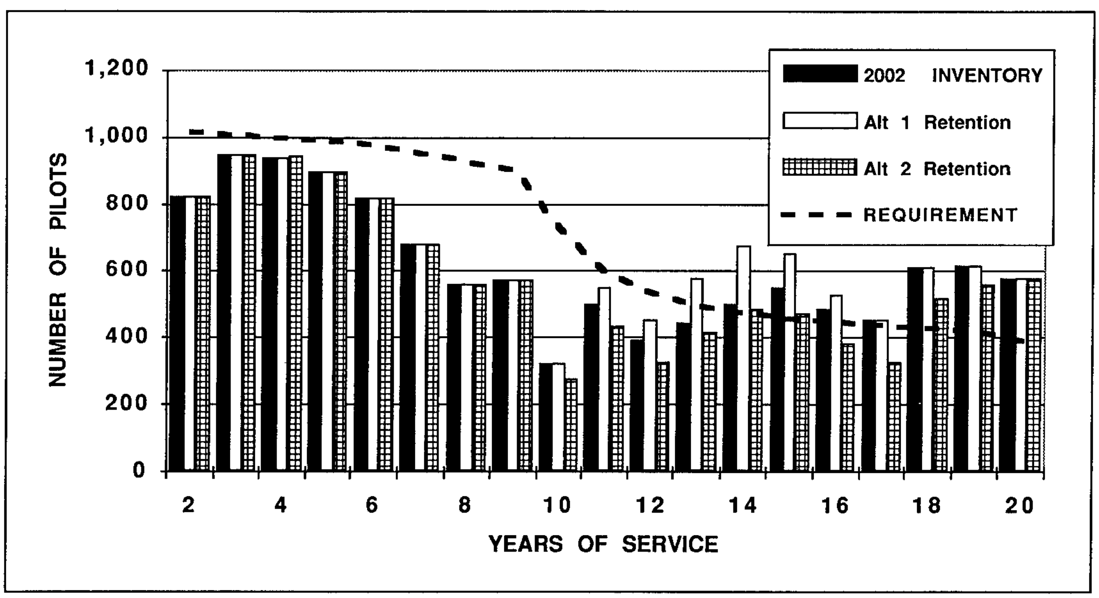


Figure 10: Effects of Retention Alternatives

revised their planned training rates upward to those shown in the second column. As a result, the base case shown does not fully mirror current Air Force estimates. We examine two alternatives: an increase in UPT of 100 starting in FY00; an increase in UPT of 100 starting in FY97 and continuing each year thereafter, as shown in Table 2. The numbers selected were premised on feasibility and reasonable access to training bases and training aircraft.[6]

Figure 11 shows the results of increased UPT relative to the base case. Increasing UPT earlier offers the best solution thus far, in terms of both increased inventory (+500 over the base case) as well as significant impacts on the bathtub. Indeed, the effects on YoS 5-7 groups is quite pronounced and the inventory profile for these groups appears to be remarkably close to the requirements line. Increasing UPT later offers a minor increase in inventory of 200 pilots, with more modest impacts on alleviating the bathtub effect.

EFFECT OF CHANGED TRAINING AND CHANGED RETENTION POLICIES

We explored one last alternative that combined both changed retention and increased training, as shown in Table 3. We thought it might be interesting to consider a real-world situation in which the shaped cohorts had lower retention, (as envisaged in Alternative 2), but that the Air Force, in response to the much lower than expected retention, increased UPT training by 100 for each year (except the first year, FY97).[7]

The result of this scenario on the FY02 inventory is presented in Figure 12. The combined effect of low retention and increased training is to increase the estimated pilot shortfall by 100. Increased training offsets to some extent the small negative effects of reduced retention because it allows the service to gain back those lost through lower retention and, in addition, mitigates the bathtub problem as well because the pilots are distributed better in terms of required experience.

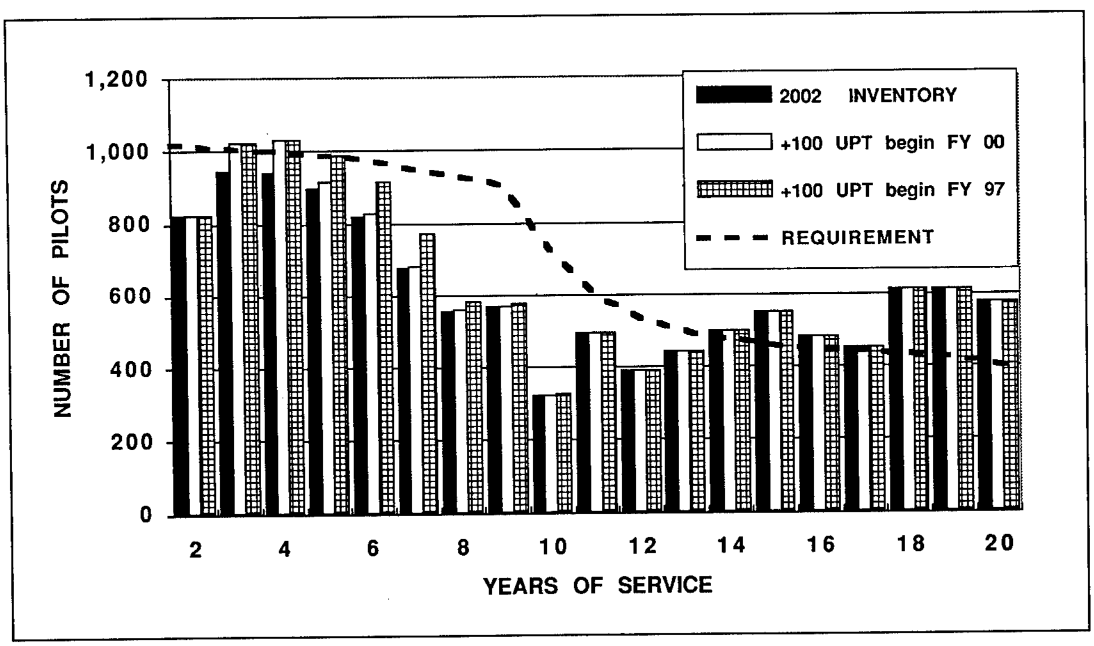


Figure 11: Effects of Training Alternatives

No likely policy scenario affecting inventory alone completely resolves the twin problems of overall shortage in the outyears and maldistribution of experience. The figure below (Figure 13) summarizes our assessment of the base case and the five scenarios that we considered: Alternative 1 (high retention), Alternative 2 (low retention), Alternative 3 (increased UPT later), Alternative 4 (increased UPT earlier), and Alternative 5 (low retention/increased UPT earlier), in terms of their effects on the predicted inventory shortfall and the bathtub. The low retention scenario offers the worst case scenario: the shortfall increases relative to the base case and, in addition, the bathtub effect worsens. Increasing UPT earlier, the policy largely adopted by the Air Force in their current plans, offers the best solution: the shortfall declines and there is a marked salutary effect on the mismatch between the required and actual experience profile. Increased retention, while it has the biggest positive effect in reducing the magnitude of the expected shortage, does not significantly affect the bathtub and increases the overage in later years of service.

CONCLUSIONS

It is evident that the Air Force is likely to face overall shortfalls of pilots by FY02 and, perhaps more serious, a significant maldistribution of experience. Changing retention does not offer much promise for changing either of these two outcomes, but reducing requirements and increasing training do. Reducing requirements even beyond the current cuts planned by the Air Force, particularly in the staff category, is crucial to reducing overall shortfalls in the future while increasing UPT is fundamental to solving the bathtub problem. However, the Air Force needs to be able to absorb graduates of UPT in cockpits and operational units and assignment policy is the key to increasing absorption. As such, a fundamental reexamination of assignment policy in order to accommodate these new goals seems warranted. Indeed, the Air Force is already doing this by moving away from the voluntary assignment system into one better suited to meet the needs of the service in terms of flows of individuals.

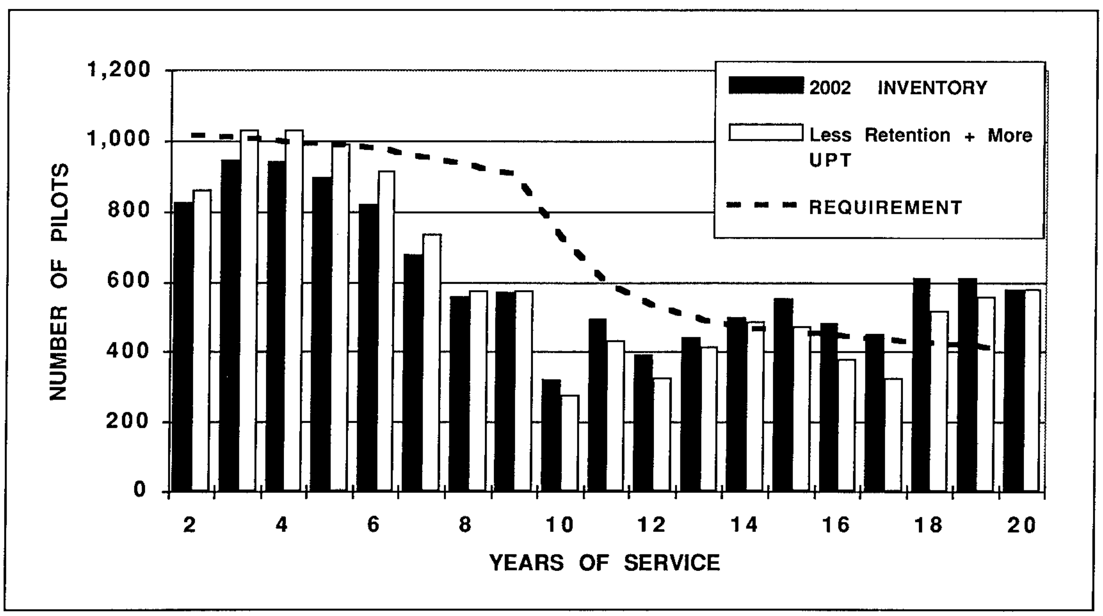


Figure 12: Effects of Lower Retention and Increased Training

However, a number of measures can be taken in the short term to minimize the effects of the bathtub and shortage problems. We list these below, without attempting any detailed assessment of them.

(1) *Active duty tours for certain ARC pilots* would directly increase active inventory. The Air Force is currently employing this policy in a modest way.[8]

(2) *Reserve/civilian instructor pilot manning* would directly decrease active-duty pilot requirements, although questions of experience and training might arise.

(3) *Changed assignment policy* would accelerate absorption of new pilots into units as more experienced pilots move into staff or other positions, opening up cockpits to new pilots. This would increase inventory.

(4) *First assignment instructor pilot (FAIP) manning* decreases MWS-qualified specific requirements because new pilots remain as instructors after finishing UPT to fill non-specific-MWS positions. In addition, this allows the FAIPs later to be absorbed into units more quickly because they have more flying hours Throughout the 1980s, the Air Force employed this policy but then dropped it in favor of using more experienced pilots as instructors. These pilots, it was felt, brought with them actual experience in operational units and were more strongly rooted in the service culture. However, resurrecting the FAIPs might be one way of decreasing requirements.

(5) *Prioritizing assignments for fill* (as has been done historically) is yet another stopgap measure that could help reduce pilot requirements and hence, expected shortages. For example, pilots could be used to fill the top-priority positions while non-rated officers or navigators could be used to fill other positions as has been done in the past or these positions could be left vacant.

ENDNOTES

1. It is important to point out that this measure is no longer considered to be a good retention measure; currently the take rate for the Aviation Continuation Pay (ACP) bonus is being used as a short term measure and Total Active Rated Service (TARS) as a long term or steady state measure to estimate future retention of pilots. However, because it was used over much of this time period, we show the CCR in the graph.

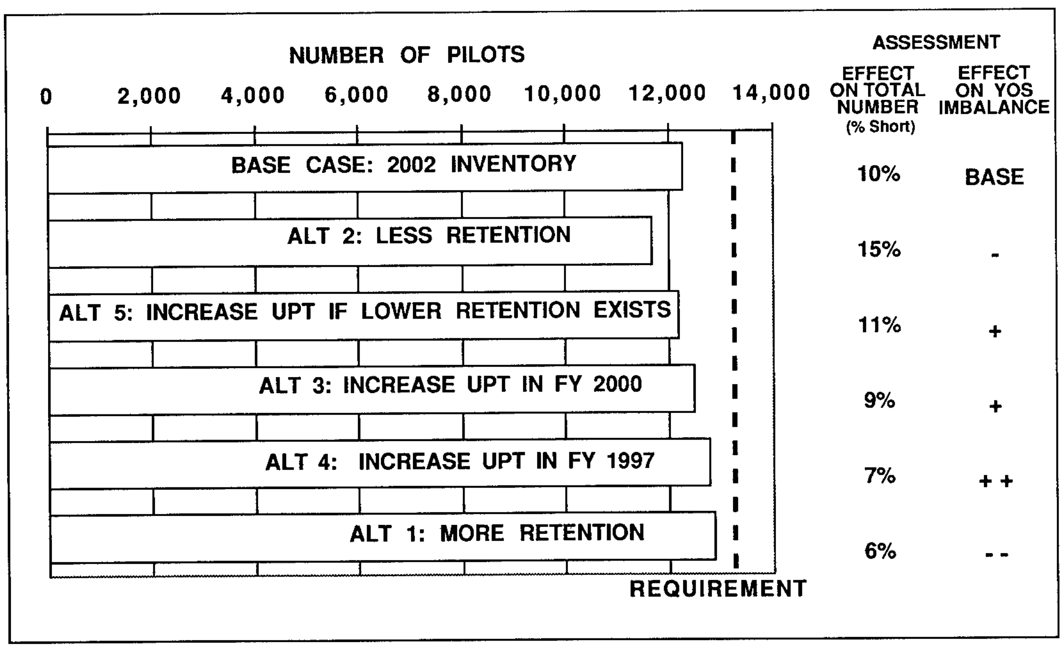


Figure 13: Assessment Summary

2. Although the Air Force has also used its Officer Training School in the past to provide a pilot training surge capability when necessary, the numbers are smaller (essentially zero for several years now), and its management of these officers differs markedly. Reserve officers on active duty in the Air Force are automatically screened for augmentation by service-wide centralized boards, while USNR officers retain a separation date and must formally apply for augmentation. The naval aviators who apply are screened separately by aircraft type for selection. Recent selection rates have been below 15 percent of those applying (for most types of aircraft).

3. Historically, the Air Force used a *sustainment model* to determine pilot production requirements, since its primary concern was to set pilot production high enough to sustain the inventory at levels that met or exceeded pilot requirements. Then in the 1970s, inventory managers recognized that this production rate coupled with reductions in absorbing cockpits had caused the proportion of experienced crewmembers to drop below minimum levels in operational units. They next adopted an *absorption model*, which they used until it became clear that PCS stability had dropped to the point (15 - 18 months for CONUS fighter units) where retention was seriously threatened. Multiple initiatives (some monetary, and others in promotion and quality of life categories) raised pilot retention to its highest levels in history in the mid-1980s, immediately before the drawdown. Then the abrupt transition into the drawdown mode, coupled with firm commitments to UPT inputs and the implementation of the Officer Voluntary Assignment System (OVAS), created pilot production overages which required pilot banking and other severe management measures. This led the Air Force to also incorporate a *vacancy driven model* (based on OVAS) in inventory management. An effort, based on this model, to eliminate the pilot bank and produce a rapid return to more traditional inventory dynamics places significant constraints on UPT production through FY 97.

4. The data shown here reflects results obtained using end-FY 93 inventory (which includes non-normal losses) as the starting point. Policy decisions were made to reduce the pilot inventory in 1992 and 1993 during the drawdown through "non-normal" losses due to Voluntary Separations Incentives/Selective Separation Bonuses, early retirement, bonus excusals, involuntary separations, and other initiatives. In later projections, we included only normal losses based on expected retention behavior. Given the inherent uncertainty of inventory projections through FY 2002, numerical results could be somewhat different. These data differences do not change our assessment. Our projection uses the current Air Force Pilot and Navigator Distribution Plan (AF/XOOT Letter, 27 September 1993) to fix UPT production and MWS distribution through FY 97 and assumes that these parameters subsequently are fixed at steady state sustainment levels.

5. Historical ranges for TARS are between 11.1 and 14.1 years.

6. During the 1980s, UPT training was typically 1,500-2,000 per year.

7. Our reasoning was that by the time the Air Force could recognize the lower retention rate and act on it, it would not be feasible to increase training by 100; hence we adpoted the lower number for FY97.

8. The Air Force recently announced a voluntary pilot-recall program for 250 fighter pilots with recent active-duty or Air Reserve Component experience.

REFERENCES

AF/XOO/DPX message 182041Z, February 1993.

Claire Mitchell Levy, *The Civilian Airline Industry's Role in Military Pilot Retention: Beggarman or Thief?* Santa Monica, CA: RAND, DB-118-OSD, 1995.

Harry J. Thie, William W. Taylor, Claire Mitchell Levy, Clifford M. Graf II, Sheila Nataraj Kirby, *A Critical Assessment of Pilot Requirements, Management, and Training*, Santa Monica, CA: RAND, DB-121-OSD, 1994.

TABLE 1

ASSUMPTIONS UNDERLYING BASE CASE AND RETENTION ALTERNATIVES

COHORT YEAR	BASE CASE	ALTERNATIVE 1: High Retention	ALTERNATIVE 2: Low Retention
FY95	13.5	13.5	12.5
FY96	13.5	13.5	12.5
FY97	13.5	13.5	12.5
FY98	12.0	13.5	12.0
FY99	12.0	13.5	12.0
FY00	12.5	13.5	12.5
FY01	12.5	13.5	12.5

TABLE 2

ASSUMPTIONS UNDERLYING BASE CASE AND TRAINING ALTERNATIVES

COHORT YEAR	BASE CASE	REVISED AF PROJECTIONS	ALTERNATIVE 3: Increase UPT Later	ALTERNATIVE 4: Increase UPT Earlier
FY95	500			
FY96	525			
FY97	670	700		770
FY98	811	925		911
FY99	911	950		1,011
FY00	990	1,025	1,090	1,090
FY01	1,050	1,050	1,150	1,150

TABLE 3

ASSUMPTIONS UNDERLYING BASE CASE AND RETENTION/TRAINING ALTERNATIVE

COHORT YEAR	BASE CASE Retention Assumptions	ALTERNATIVE 5: Low Retention	BASE CASE Training Assumptions	ALTERNATIVE 5: Increase UPT (After Lower Retention Is Known)
FY95	13.5	12.5	500	
FY96	13.5	12.5	525	
FY97	13.5	12.5	670	720
FY98	12.0	12.0	811	911
FY99	12.0	12.0	911	1,011
FY00	12.5	12.5	990	1,090
FY01	12.5	12.5	1,050	1,150

Modelling the Mobile Land Battle: The Lanchester Frame of Reference and Some Key Issues at the Tactical Level

L. R. Speight

Cranfield University, Royal Military College of Science, Shrivenham, England.

INTRODUCTION

It is now some eight decades since Lanchester (1916) and Osipov (1915), probably independently, put forward their theories of battle attrition in terms of a system of simple coupled differential equations. Since that time this scheme, embellished of course to accommodate much more complex assumptions than were originally considered, has become the standard frame of reference for much battle modelling. To be sure, computer technology has also advanced tremendously during this time, and so these summary methods have been complemented by more literal, and very detailed, representations of the processes of war. But still Lanchester formulations have a pervasive influence as conceptual tools; as the actual building blocks of large scale simulations; and as a means of interpolating between, or even extrapolating from, the results obtained from more detailed investigations.

Detailed criticisms of Lanchester theory have been produced. Many attempts have also been made to compare its predictions with the actual outcomes of historical battles. But it is not the intention of this paper to criticize Lanchester theory, except perhaps obliquely when considering the features of battle which any summary method, Lanchester or otherwise, ought to take into account. Still less is it the intention to attempt any sort of historical validation. Indeed, it is a central tenet of this paper that there are severe limitations on what can be achieved simply by comparing the predictions of a complete theory directly with the outcomes of trials or of real battles. Helmbold has set out very clearly some of the difficulties in his recent series of articles in *Phalanx* (Helmbold, 1993). In most cases the historical data are limited to two points: the starting and the (sometimes dubious) terminal strengths of the opposing forces, giving very little scope for the fitting of any model parameters. But, above all for such a variable and complex phenomenon as battle, the number of actual instances is not large, each instance being affected by particular sets of circumstances which might be expected to have a major impact on the outcome. There seems to be a case for an approach which could complement that of direct comparison, which would invoke a set of intermediate constructs or intervening variables, acting as a bridge between detailed investigations and grand theory. Such intermediate hypotheses could be validated by appeal to evidence of a different kind than just that of overall battle outcomes. When linked to form a coherent account of battle dynamics, this latter could be exercised in a systematic manner in the controlled environment of the computer laboratory. In this way our understanding should be advanced by the usual processes of science: constant query and debate, the generation and testing of hypotheses, and by focusing attention and effort on areas which seem to be critical for model predictions. But in the meantime grand theory which flies in the face of the evidence collected in this way should at least be queried: any overall model must be consistent with the mechanisms which implicitly underpin it.

This paper concentrates on the direct fire armour anti-armour battle at the tactical level. Since the aim is to establish basic principles, rather than to attempt practical predictions, discussion is restricted to 'homogeneous' forces. The document therefore starts with an extremely brief account of Lanchester theory, bringing out only those few points germane to the arguments which follow. The main portion of the paper sets out some modelling issues under six different subheadings. Four of these have to do with the logical components of any battle model at this level, but the stress is on the pattern of dependencies within and between them. This discussion leans heavily on the evidence from field trials and, to a lesser extent, from war records. Another subheading has to do with the universally observed phenomenon of the degradation of combat skills under live fire. The last alludes to the class of model to which the Lanchester formulation belongs. Only the issues raised under the final subheading are addressed further in this paper, and then only in sufficient depth to set the scene for two forthcoming papers, which later will take up the remaining modelling points in more detail.

THE LANCHESTER FRAME OF REFERENCE

Full accounts of the many developments of Lanchester theory and practice have been given elsewhere, notably in the compendious monographs written by Taylor (1980,

1983). For present purposes it is only necessary to set out the most fundamental details. Table 1 lists some simple alternative formulations. Here x and y represent the numbers of the X and Y forces at any given time, and a and b are positive coefficients, which do not depend on time, reflecting the ability of one side to inflict casualties on the other.

Table 1. Lanchester formulations.

Law	Attrition equations
Lanchester 'Square' Law for 'aimed' fire	$dx/dt = -ay$ $dy/dt = -bx$
Lanchester 'Linear' Law for 'unaimed' or 'area' fire	$dx/dt = -axy$ $dy/dt = -bxy$
Lanchester 'Mixed' Law (Brackney, 1959)	$dx/dt = -ay$ $dy/dt = -bxy$
Lanchester 'Power' Law (Helmbold, 1965)	$dx/dt = -ay(x/y)^c$ $dx/dt = -bx(y/x)^c$

Lanchester's 'Square' Law is consistent with the following main assumptions:

a. The forces use 'aimed' fire, so that the location of each target is known in advance, the effect of each firing is apparent, and fire is shifted to a fresh live target as soon as a kill is achieved. The time taken to shift fire does not depend on the number of live targets remaining.
b. The forces are within weapon range of each other.
c. The effect of each weapon round is independent of the effect of other rounds.
d. Fire is uniformly distributed over all live enemy targets.
e. Each force is 'homogeneous' (that is to say it consists of one weapon type, or of types with identical lethalities and vulnerabilities).
f. All forces are committed at the beginning of the battle, and there are no reinforcements.

The conditions listed under assumption (a) above are those which Lanchester believed were appropriate to 'modern' warfare, with a high degree of weapon control and precision. Target acquisition times and kill rates might then be regarded as independent of the number of live enemy weapons remaining. Where these conditions do not obtain then the assailant must resort to 'area' or 'unaimed' fire, and kill rates might be expected to vary in proportion to the surviving target density. Various proposals have been put forward to modify these two alternative assumptions. Thus Brackney (1959) has suggested a 'mixed' law on the basis that, in general, the defender has little difficulty in acquiring targets, whereas search times for the attacker may become proportionately longer as the numbers of defenders decrease. Helmbold (1965) argued, on the basis of much historical evidence, that as the ratio of force levels of two opposing sides increased, so the side with the larger numbers would have increasing difficulty in bringing all its combat potential to bear. Sheer limitations of available space, to say nothing of terrain masking and reaction time effects, might limit the effectiveness of the stronger side. He therefore proposed that the Lanchester 'square' law should be modified by adding a (symmetrical) power function of the surviving force ratios. Other Lanchester laws are listed by Helmbold (1993) in the *Phalanx* discussion articles already referred to. One feature that these different formulations have in common is that kill rates and vulnerabilities are all concentrated in the a and b coefficients, and the effects of opposing numbers are concentrated in the x and y terms of the equations. Implicitly, a is the probability density per unit of time that y will kill x in a one-on-one duel, and b is the reverse probability density (although few, perhaps, would wish to push the logic of the model thus far).

Obviously, the strict assumptions underlying Lanchester's 'Square' Law cannot possibly hold true for units or formations of even a moderate size, especially assumption (b). Nevertheless, Lanchester formulations are very often used at an aggregated level to calculate the expected attrition when two formations oppose each other. At this level Lanchester can hardly be characterized as a theory, but more as a simple descriptive device or belief. For Lanchester's 'Square' Law the description might go roughly as follows: "At this level of warfare a reasonable approximation of the expected attrition process may be obtained by supposing that each surviving weapon system of force X in the 'contact zone' (however that may be defined), or in

different parcels of the 'contact zone' (however they may be defined), will on average kill surviving weapon systems of force Y in that zone (or parcel) at a constant rate b. Similarly, each surviving weapon system of force Y in the 'contact zone', or in different parcels of the 'contact zone', will on average kill surviving weapon systems of force X in that zone (or parcel) at a constant rate a." The actual attrition rates will, of course, be a function of the timing and pattern of encounter between groups of weapon systems on each side, as well as of the attrition following such encounters, which will in turn be conditioned by the local circumstances of each encounter. The course of reinforcement to the 'contact zone' is then an important consideration. Most of the historical validation of Lanchester theory has taken place at the operational or strategic, rather than the tactical level. Despite the majority finding that neither the 'Square' nor the 'Linear' Law fits the historical results very well (see, e.g., Osipov, 1915; Helmbold, 1961a, 1961b, 1964; Willard, 1962; Weiss, 1966; and Hartley, 1991), 'Square' Law formulations remain popular as the basis for attrition calculations for direct fire weapons at an aggregated level, and 'Linear' Law formulations as the basis for the attrition due to indirect fire weapons.

As the level of aggregation is lowered there must surely come a point where the strict conditions underpinning Lanchester theory could potentially hold true. In fact, for the direct fire battle at least, this point is likely to be down at the mini-battle level, with only a small handful of weapon systems involved. The paradox is that, by the time we reach the level of the small unit engagement, the need for considerable embellishment of the basic Lanchester formulations becomes inescapable if the latter are to have any real descriptive power. The opening up of lines of sight, rates of target acquisition, kill probabilities given a firing, and in most cases the balance of advantage will all shift as opposing forces close and the engagement progresses. The particular terrain, tactics and visibility conditions will all have profound and interacting effects. Direct validation at this level, to put it no stronger, becomes extremely difficult. Hence the belief, already stated, that intervening constructs and variables should be correctly identified, and that a controlled, quasi-laboratory, experimental testing environment is desirable as an adjunct to direct validation.

It is for the reasons just discussed that we have referred to the 'Lanchester frame of reference', rather than to 'Lanchester theory' in the title of this paper and in the heading of this section. There seems to be little prospect that Lanchester theory will serve as a valid description of the attrition dynamics of real battles, except in an approximate, rough and ready, manner. Ancker and Gafarian (1988) have examined this matter from a mathematical point of view. They have shown that the deterministic differential equations of the Lanchester frame of reference will, in general, provide poor approximations to the results yielded by their stochastic equivalents. Furthermore, the validity of a stochastic Lanchester formulation as a representation of the attrition process rests on fairly restrictive assumptions as to the properties of the underlying interfiring time distributions. Nevertheless, these kinds of result seem likely to do little to lessen the attachment of model builders to Lanchester formulations, at least until an alternative frame of reference presents itself, as mathematically convenient and based on assumptions which are evidently more realistic.

Rather than starting from some theoretical framework, the approach taken in this series of articles will be to examine the evidence from live trials and from war, and then try and find answers to questions of the following kind. What mechanisms or processes appear to be the most important in affecting the course of live battle? What are their properties in a statistical sense, and how do they interact? When we have combined these elements into a coherent description of the dynamics of a mobile land engagement, what kind of relationships do they then yield between assumed battle conditions or parameters and battle outcomes? Can we find a Lanchester (or other) formulation which will provide a reasonable approximation to the predictions based on this logic and evidence, and if so which Lanchester (or other) variant is preferable? As we have already suggested, it is not really possible to answer a question such as the last by appealing directly to battle outcomes. With the limited degrees of freedom provided by just the sample statistics of starting and terminal

force strengths one is, in effect, reduced to deciding some simple form of relationship in advance, and then using the data to estimate the relevant parameters. We need additional evidence if we are to place our modelling practice on a firmer logical footing.

SOME MODELLING ISSUES

Any battle model at the tactical level must include the following components:

a. A statistical description, at least, of the opportunities for elements of the two forces to engage.
b. A portrayal of the target acquisition process, given that such opportunities occur.
c. A representation of the aiming process and the computation of the chances of a kill, given that targets are acquired to this pattern.
d. A logic for deciding that the battle is effectively over, given that it unfolds in this manner.

If the complete model is to have relevance to wartime conditions it must also reflect what is known of:

e. Combat degradation.

At various stages in our programme of enquiry we should also pause to consider:

f. The likely utility of the Lanchester frame of reference at this level of aggregation.

Opportunities for Engagement

All the evidence of trials and war records suggests that tactics, terrain and manoeuvre combine to split an engagement between two different formations into a series of parallel and/or sequential 'mini-engagements', with very few combatants participating in each. Figure 1 illustrates this phenomenon.

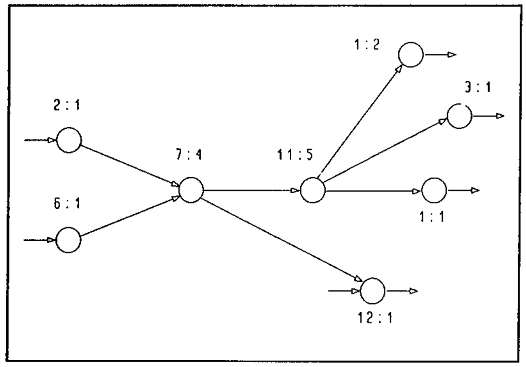


Figure 1. Representation of a succession of mini-battles from a field exercise. Each circle corresponds to one such mini-battle, with the numbers representing the participating attack: defense weapons. The arrows indicate the flow of some weapons from one engagement to another. (Adapted, with permission, from Figure 3 of Rowland, 1984).

In the series of field trials from which this illustration was taken the odds in these mini-battles seemed to vary in lognormal fashion about the local average, the mean numbers of participants being only some 2.5 for the defenders and 5.5 for the attackers. In the conditions of World War II, also, fire fights typically had only single numbers of participants on each side.

The structuring of mobile engagements into a series of mini-battles is of importance for three reasons. Firstly, the formation of mini-battles will set a natural upper bound on the intensity of the putative engagement as a whole and on the pace with which it can unfold. Secondly, even if each mini-battle adhered closely to one or other of the Lanchester formulations, then the outcome of an ensemble of mini-battles would almost certainly differ from that of a single Lanchester battle between the two complete sets of combatants (see, for example, Ancker and Gafarian, 1988, and McNaught, 1995). And, lastly but most importantly, the conditions of each mini-battle (especially its time span and the numbers in each opposing subset of combatants) will have a fundamental effect on the whole process of target acquisition.

Target Acquisition given the Opportunity to Engage

The probabilities of who detects who, in what sequence, are far more important determinants of the pattern of attrition than are kill mechanisms, given that a target has been acquired. Search is a complex topic in its own right, and reliable models of visual performance are at least as elaborate as most battle models. The ORACLE model (see, e.g., Cooke, Stanley and Hinton, 1994; Hinton, 1992; and Overington, 1982) is an example of this kind, validated against a host of field trials and laboratory experiments, and taking account of the various characteristics of the target; its background; the atmosphere; the observer, together with any sensor or display; etc.; etc. Without going into all these complexities, it seems sensible to summarize a few broad trends which will affect the dynamics of any exchange of fire.

During search the eye does not scan smoothly, but by rapid intermittent movements, interspersed with stationary dwell times of roughly one third to one half of a second. If a target is detected during such an interval the eye then moves to fixate it in the centre of the field of view. Most visual performance models therefore use the characteristics alluded to in the previous paragraph to compute a 'single glimpse' visual lobe: a function which will predict the (decreasing) probability of detection as the target is moved from the temporary centre of fixation to the periphery of the field of view. A 'conspicuous' target will have a 'large' visual lobe, so that the target is likely to be detected almost wherever it is imaged on the retina. An 'inconspicuous' target will have to be imaged near the centre of fixation to have any likelihood of being detected; and, assuming that the eye moves in quasi-random fashion from one fixation to another, the 'single glimpse' probability of detection will be roughly the ratio of the (suitably integrated) visual lobe area to that of the search arc. Once it is detected it must be recognized (placed correctly within a category of possible target types) and then identified (given the correct label within that category). Experiments and trials suggest that, to a first approximation, the detection probabilities for successive glimpses can often be regarded as independent, so that a geometric distribution will adequately describe detection times in discrete time models, and an exponential distribution will serve a continuous time model. From this we can see that the size of the visual lobe, the width of the search arc, and the spacing of the potential targets will all be important in determining the probable time history of acquisitions. To the extent that search arcs are small, and targets are 'conspicuous' and closely grouped (more likely to be true of the attacker as seen by the defender), so will mean acquisition times tend to be independent of the number of targets. To the extent that required search arcs are large, and targets are 'inconspicuous' and dispersed (generally typical of the defender seen from the vantage point of the attacker), so will mean acquisition times tend to be inversely proportional to the number of targets.

There are some obvious implications in all this for the manner in which target acquisition should be dealt with in any battle model. Manifestly, an assailant can only detect those direct fire targets which are actually present in the field of view. Targets which are hidden by terrain screening or other forms of obscuration cannot be candidates for detection, and still less can those in the opposing force at large but remote from the viewer. But, given the opportunity of two subsets of the opposing forces to interact, the scene presented to the viewer is crucial. The 'per glimpse' chances of detection will depend, *inter alia*, on the number of opponents actually present, but also on their spacing and on the width of the search arc. If targets are closely spaced compared to the width of the visual lobe, then their chances of detection will not be independent of each other: to miss one will be, by and large, to miss its neighbour. As the range of engagement decreases so will the visual task alter, as the angular subtenses and dispersions of targets increase and as the atmospheric attenuation decreases. In short, given the potential for a mini-engagement to occur, target acquisition times are likely to be somewhat complex functions of the numbers of opponents involved and of a number of other factors, including range. Only in extreme circumstances will this function approximate to the assumptions inherent in the Lanchester 'Square' Law. However, this approxi-

mation is likely to be closer for the defending force than for the attacking one.

A further major complicating factor is the effect of firing signatures on the target acquisition process. Compared to the visual signatures of the defending targets themselves, their firing signatures tend to be extremely conspicuous. This is confirmed by the findings of trials, where the probability of a defending weapon system being killed seems to be in almost direct proportion to the number of times it fires. This means that the target acquisition rate of the attackers is likely to be determined in large part by the firing rate of the defenders, which in turn is likely to be a function both of the number of defenders and (through the detection process) of the attackers.

'Overkill' is another phenomenon which seems to be all-pervasive in trials and in war. There are two main mechanisms by which overkill can arise. Firstly, under the pressures of live conditions there may be some difficulty in discriminating between 'live' and 'dead' targets. Secondly, irrespective of this difficulty, there is likely to be a sub-optimal distribution of engagement effort. This can be seen most clearly in the case of attackers where, as just discussed, the principal detection cue is typically the defender's firing signature. The first defender to fire is likely to attract a disproportionate amount of the attacker's return fire (especially as it is hard for the latter to take any account of defenders who are still undetected). The second defender to fire is likely to attract less of the attacker's effort, if only because of the numbers now committed to the engagement of the first firer; and so on. Important, too, is not just the distribution of effort but its sequencing: by the time an assailant has started to prepare his response to a firing signature the associated munition is already in flight.

In practice fire coordination may further complicate the picture. Defending field commanders will generally attempt to trade on their advantage in target acquisition, holding fire until the attacker is committed and covered by several defense weapons, before disclosing their own positions by the telltale act of firing. It can be difficult to incorporate these effects of judgement, calculation and prediction in any purely statistical model. Defenders may also jockey after only a few firings in order to reduce their chances of being pinpointed by the opposition.

The conclusion from this subsection of the present paper must be that target acquisition rates, and the distribution of engagement effort, must be a complex function of the numbers of the opposing forces interacting within any one mini-battle, affected by visual detection dynamics, firing signature characteristics, overkill and by the command and control process. But, it is worth emphasizing once again, it is the mini-battle which provides the context for these target acquisition effects.

Kill Probabilities given Target Acquisition

Even in times of peace, research and development into the ways and means of achieving military destruction continues apace. Because it is at the cutting edge of technology this topic is subject to a great deal of assessment and evaluation. This is not to say that all the problems of computing kill probabilities have been solved. But compared to the other elements discussed here this matter has received a considerable amount of attention, and so it will be dwelt on no further in this paper.

Battle Termination Criteria

Live trials, let alone actual battles, can be messy and confusing affairs, full of minor clashes, chance encounters, penetrations without contact and the like. It is natural, though, that modellers should concentrate on well-structured (by implication intentional and possibly key) engagements, which have a definite beginning and an end. But the end will depend in large part on the mission and goals (and, for the loser, the contingency plans) of the field commanders. Occasionally these missions can be couched in terms of the destruction of enemy forces. Occasionally, too, they can concentrate neither on destruction nor on territorial gain, but on tasks such as distracting or 'fixing' opposing formations. But more often military objectives are couched in terms of time and space. Manoeuvre is the key to the armoured battle.

Lanchester theory is, of course, concerned with attrition, and Lanchester formulations include the dimension of time but do not explicitly include the dimensions of space. As a consequence modellers have tended to concentrate on attrition considerations, such as so-called 'breakpoint' thresholds, in determining the criteria for battle termination. This despite the fact military history does not provide clear guidelines as to what these 'breakpoint' levels of attrition may be. However, in many cases it seems that the battle effectively ends, not on account of the casualty levels sustained by one side or the other, but because one commander has achieved his objective and the other has recognized this fact. It is perhaps worth emphasizing at this point that attrition is normally but one means to an end: the imposition of one side's will on the opponent. As Sun-Tzu wrote more than two thousand years ago: "To fight and conquer in all your battles is not supreme excellence; supreme excellence consists in breaking the enemy's resistance without fighting."

Space, although not included explicitly, is often included implicitly in most elaborate Lanchester formulations as a function of time. But what so often happens in modelling approaches which are careless in preserving the interactions and dependencies between the processes of war is that, even if the relativities between the two opposing forces are preserved, the overall rate of attrition becomes distorted and unrealistic. The implicit relationship with space is then damaged. This effect can be exacerbated by the widespread degradation of military skills, and hence reduced kill rates, normally experienced in battle.

Combat Degradation

There is every possibility that coordination and judgemental skills may be degraded in battle conditions, given that they are exercised in the first place. There is ample evidence from the psychological laboratory and from trials that too high (or too low) a level of arousal reduces the level of performance, and there could hardly be a more arousing environment than that of live battle. However, obtaining reasonable estimates of the likely level of degradation for each and every military task or task component raises severe problems, both methodological and because of the sheer extent of the effort required. Reluctantly, one must conclude that the accumulation of knowledge on this topic will be gradual and piecemeal, and that at present it is hardly possible to assemble existing information into a coherent body of battle performance estimates.

However, there is another form of combat degradation which will be considered in this series of papers. This is not the exercise of military skills at a reduced level of effectiveness, but the refusal or inability to exercise them at all. The classical study of this aspect of war was that of Marshall (1947), based on extensive post-operational interviews of U.S. troops. This study, suggested that a very significant proportion of those nominally involved made no effective contribution to the battle whatsoever. If individuals do differ very significantly in the extent to which they participate in the fight then it would be most unwise to factor this in as, say, a simple reduction of kill rates. Proper account must be taken, not just of the likely level of participation, but of its variability from individual to individual or from one weapon system to another.

The Likely Utility of the Lanchester Frame of Reference

The limitations of the Lanchester frame of reference have already been discussed, particularly in its deterministic and continuous force level form. However, science is replete with examples where continuous and deterministic models serve well as descriptors of the aggregated effects of quite complicated, discrete and stochastic phenomena. Having attempted to tease out the most important mechanisms of the mobile battle, and tried to describe the way in which they interact, it seems sensible to pause and evaluate Lanchester theory once more. Does some Lanchester formulation provide a reasonable summary description of our more detailed results, and is the continuous and deterministic form likely to introduce unacceptable bias compared to one which is discrete and stochastic?

LANCHESTER AS A FRAME OF REFERENCE FOR THE MINI-BATTLE

A number of writers have sought to investigate the statistical properties of the stochastic Lanchester formulation (see, for example, Brown, 1963; Goldie, 1977; Smith, 1965; Watson, 1976; Weale, 1992; and Weiss, 1963). However, these results are not really useful in the context of the mini-battle. Either they restrict themselves to one or other of the 'Square' or 'Linear' Laws; or they look to large sample approximations to reduce the amount of computation required. To provide a preliminary answer to the two questions posed at the end of the last section we have used a discrete time Markov model devised previously (Speight, 1995), which provides a very simple stochastic and discrete force equivalent to the deterministic and continuous Lanchester formulation. In general terms the latter can be stated thus:

$$\begin{aligned} dx/dt &= af_1(x,y) \\ dy/dt &= bf_2(x,y) \end{aligned} \qquad (1)$$

where $f_1(x,y)=y$ for the 'Square' Law;
$f_1(x,y) \;=xy$ for the 'Linear' Law; and so on.

In the Markov equivalent let S be a state matrix with dimensions $x=0, 1, \ldots, x_0$; $y=0, 1, \ldots, y_0$, where x_0 and y_0 are the numbers of the X and Y assailants present at the start of a mini-battle. At the outset, then, the probability of being in state $S(x_0,y_0)$ is 1, and that of being in all other states is zero. If the system is in state $S(x,y)$ at some time t, then the probability of a transition to the same or other states at time $t+\delta t$ is held to be some function of x and y. Let us then define two functions, $g_1(x,y)$ and $g_2(x,y)$. In the model which we have used the transition probabilities are given by:

For: $x=1,\ldots,x_0$
$y=1,\ldots,y_0$

$$\left.\begin{aligned} &P(x-1,y \mid x,y)=\alpha g_1(x,y)\{1-\beta g_2(x,y)\} \\ &P(x,y-1 \mid x,y)=\beta g_2(x,y)\{1-\alpha g_1(x,y)\} \\ &P(x-1,y-1 \mid x,y)=\alpha g_1(x,y)\beta g_2(x,y) \\ &P(x,y \mid x,y)=\{1-\alpha g_1(x,y)\}\{1-\beta g_2(x,y)\} \end{aligned}\right] \qquad (2)$$

$$P(x,0 \mid x,0)=1 \qquad x=0,\ldots,x_0$$

$$P(0,y \mid 0,y)=1 \qquad y=0,\ldots,y_0$$

where $P(x',y' \mid x,y)$ is the conditional probability of moving from state (x,y) to (x',y') in one step.

The transition probabilities between all pairs of states not specified above are zero. If the time interval δt in this formulation is made very small it will approximate to a continuous time stochastic Lanchester model. The parallels between this formulation and the generalized Lanchester equations (1) will be apparent.

Lanchester versus Markov Comparisons at the Mini-Battle Level

Figures 2 and 3 illustrate some of the differences between Lanchester and Markov predictions when numbers engaged are small. The first two diagrams in each figure show the Lanchester x and y time traces, as well as the corresponding expected values of the x and y marginal distributions for the Markov simulation. Starting values for the first diagram were $x_0=5$ and $y_0=10$, and for the second they were $x_0=2$ and $y_0=4$. We have picked a case where Lanchester and Markov predictions differ quite markedly, with the ratio of the a and b attrition coefficients chosen to give a 'balanced' battle (with the x and y values of the Lanchester outputs both reaching zero at time infinity). In each case the simulation was halted at the time when Lanchester attrition levels reached 75% of the starting values. The last diagram in each figure shows the probabilities in each cell of the Markov state matrix at the time that the simulation was stopped.

It will be seen that the expected number predictions of the Lanchester and Markov formulations diverge as time progresses, the Markov model predicting lower attrition in each case. The divergence is worse for the 'Square' Law than for the 'Linear' Law, and proportionately worse for 'very small', as opposed to 'small', numbers.

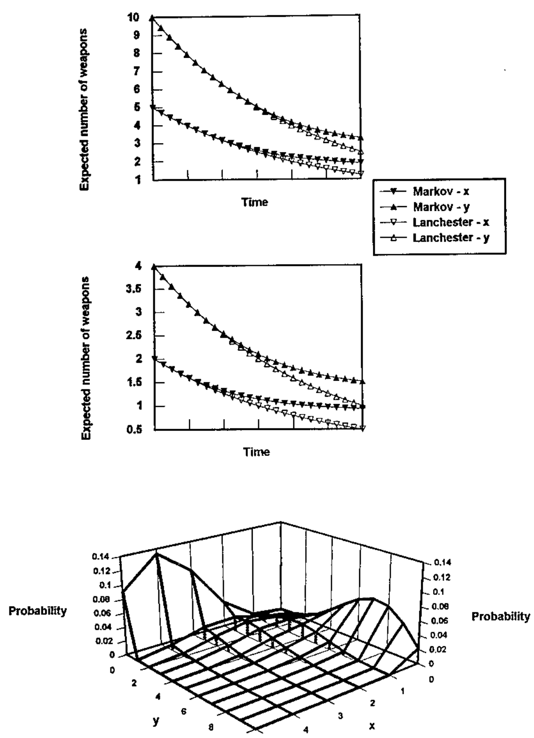


Figure 2. 'SQUARE' LAW - Lanchester versus Markov comparison. The top two diagrams show the expected value plots with time for two different starting sets of X and Y forces. The bottom diagram shows the bivariate plot of the Markov state matrix at the end point of the simulation.

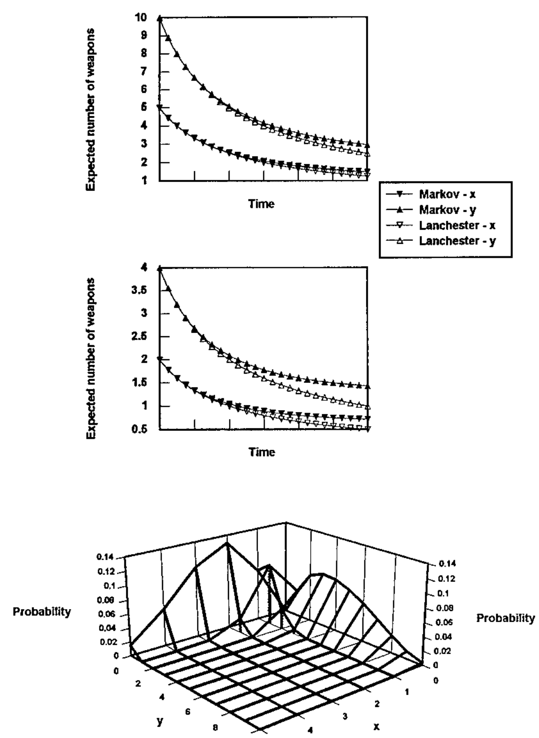


Figure 3. 'LINEAR' LAW Lanchester versus Markov comparison. The top two diagrams show the expected value plots with time for two different starting sets of X and Y forces. The bottom diagram shows the bivariate plot of the Markov state matrix at the end point of the simulation.

The bivariate plots show that the expected value trends mask very different underlying characteristics for the two formulations. By the time that the simulation was halted the Markov (x,y) distribution had virtually split into two. For the 'Square' Law one subset of the distribution had zero y remaining, with an x mode at 4 of the original 5 weapons; another subset having zero x, with a y mode at 7 of the original 10. It is perhaps a sobering thought that the single Lanchester (x,y) point at this time is quite remote from the main concentrations of probability. It will be noted that the Markov 'Linear' Law process is not so divergent as that for the 'Square' Law: this is due to the feed-back effect in the former of each side's attrition diminishing as its numbers fall.

Having made these points, it would be easy to over-dramatize the likely consequences of these sorts of differences in the context of the most widely used approaches to battle modelling. Most aggregated models are based on deterministic (hopefully expected value) attrition equations. But, as already mentioned, the evidence from trials and from war suggests that very few shots are fired within most mini-battles, and that normally the resulting level of attrition is low. Under these conditions Lanchester and Markov predictions do not diverge markedly. Perhaps the error introduced from this source is small compared to some of the other approximations embedded in most practical models, especially if (as has already been argued) the conditions assumed for the 'Square' Law are an ideal which will very rarely be attained in practice. Provided that the analyst is concerned to construct an expected value model, and (a very big 'and') to ignore the considerable variance (and at times oddly shaped distributions) about these expectations, the choice between continuous-and-deterministic or discrete-and-stochastic modelling approaches does not seem likely to be the most important issue. Far more important is to be assured that a Lanchester formulation, or a Markov equivalent, can indeed reflect in an adequate fashion the interactions and dependencies within a representative mini-battle.

Markov 'Lanchester Equivalent' Fits to a Mini-Battle Simulation

A first step in seeking the assurance mentioned in the last paragraph was to build a very simple time-stepped, state space based, Monte Carlo simulation of a mini-battle, incorporating some of the features which seemed to be of importance from our review of the evidence from trials and from war. An attempt was then made to emulate the results, using a version of the Markov model given above.

The intention when building the simulation was not to represent all the complexities of a mini-battle as they might occur in real life, but in just sufficient detail to preserve the main interactions between attackers and defenders. The set of possible states for the defenders were (Jockey, Search 1, Lay 1, Observe Shot 1, Search 2, Lay 2, Observe Shot 2, Destroyed); and those for the attackers were (Search, Lay, Observe Shot, Killed). Thus, a defender would fire a maximum of two shots before jockeying, and was assumed to be invisible and invulnerable while carrying out this manoeuvre. Assuming that a firer was not killed in the meantime, the Observe Shot state was maintained for a fixed number of time intervals, equivalent to the time of flight of the round or missile. The Jockey and Lay states were maintained for a fixed number of time intervals, with the probability of exiting from the state being a constant for each interval thereafter.

Both attackers and defenders were held to detect only targets which were not already killed or destroyed. The defenders were assumed to have a fixed high (0.5) 'single glimpse' probability of detecting a live attacker. The attackers were assumed to have to have a fixed low (0.01) 'single glimpse' probability of detecting a live non-jockeying target, provided that it had not fired during that time interval. If a defender had fired, then the probability of each attacker detecting the firer was high (0.5) for that interval only. If a detection occurred, and there was more than one target in the state associated with that detection, then the target actually acquired from that subset was decided at random. If an acquired target was killed or destroyed by another firer while the engager was in the Lay or Observe Shot states, then the engager would move on to the

next Search or Jockey states, as appropriate. The assumed kill probabilities, given that all the other steps in the engagement sequence had been completed successfully, were markedly higher for the defender than for the attacker (0.2 as opposed to 0.075).

In this simulation we have watered down the evidence on overkill, and have represented target acquisition in a way that should bring it as much in line as possible with the underlying assumptions for Lanchester's 'Square' Law. The combatants do not acquire targets which are already dead. Once one has been acquired they are assumed to have perfect and instantaneous knowledge if it is then killed by another agent, and the only penalty is then to lose the time already spent on an aborted engagement. 'Single glimpse' detection probabilities are not held to increase as the numbers of live targets increase, nor as the numbers of firing signatures within a time interval increase.

Turning to the Markov model to be fitted to the results from this simulation, various alternatives were considered for the functions g_1 and g_2 in equations (2) above. Eventually, after testing some of these alternatives to see which gave the best fits to simulation outputs in the least squares sense, the following forms were selected:

$$g_1(x,y)=x^T y^U; \qquad g_2(x,y)=x^V y^W \tag{3}$$

Lanchester's 'Square' and 'Linear' Laws, Brackney's mixed law and Helmbold's modified Lanchester would all be special cases of these formulae.

In order to fit the Markov model parameters to the results the simulation was first run as a one-on-one stochastic duel. In 25,000 runs of the simulation the proportion of each of the two assailants surviving was noted at 15 different time intervals, arranged according to a Fibonacci sequence (roughly evenly spaced on a logarithmic scale). A nonlinear least squares routine was then run to determine the values of α and β in equation (2) above giving the best fit to these proportions. A visual plot of the results is shown in Figure 4. It will be seen that the fit is not perfect for small values of t, although it improves as t increases. Some may conclude that a model of this kind provides a good enough approximation to the simulation results for most practical purposes.

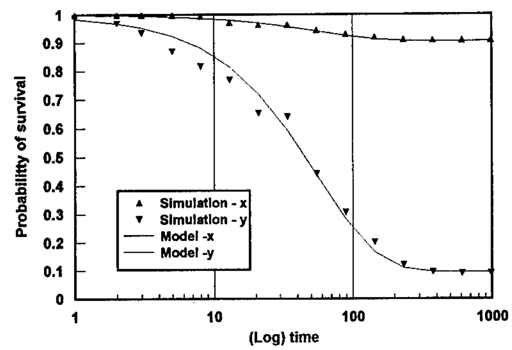


Figure 4. Markov model fit to simulation results for one-on-one duel.

The simulation was then run for a number of different x_0 and y_0 inputs. Based on 10,000 simulation runs, and using the α and β values estimated from the one-on-one duel, the values of T, U, V and W were determined which gave the best fit to the observed proportions in the (x_0+1) by (y_0+1) cells of the state matrices at the 15 sampled time intervals. Although not shown here, visual inspection of the plotted results showed the same trends as for the one-on-one duel: a reasonable fit overall, but with some discrepancies for small t. The best fit values for the parameters in the Markov model are given in Table 2.

The obvious difficulty with these results is that the best fit values of the stochastic Lanchester parameters differ markedly as a function of the starting numbers of the mini-battle. Nor is there a straightforward trend in the fitted values as x_0 and y_0 increase, and so a simple modification of the model would seem unlikely to lead to a better fit. It should be remembered that the changes of assumptions from those implicit in the Lanchester 'Square' Law have been quite minor: some lost time due to 'overkill' (but with no failure to discriminate between live and dead targets) and intermittent jockeying by live defenders, but with a significant influence on target acquisition of firing signatures. It seems quite possible that even a stochastic Lanchester formulation, separating as it does the values of the attrition coefficients from the effects of opposing numbers, may have some diffi-

culty in properly reflecting the dynamics and interactions of a putative mini-battle.

Table 2. Fitted values of T, U, V and W as a function of simulation x_0 and y_0 inputs. SSQ denotes the residual sum of squares. The values of α and β used ($x10^3$) were 0.169 and 1.596 respectively (for which the SSQ was 0.0386).

y_0		x_0 = 2	x_0 = 4	x_0 = 8
	T	0.23	0.28	0.24
	U	1.31	1.28	1.34
2	V	0.38	0.59	0.74
	W	0.87	1.24	1.20
	SSQ	0.0557	0.0468	0.0525
	T	0.00	0.40	0.49
	U	1.21	0.92	0.49
4	V	0.25	0.36	0.52
	W	0.38	0.78	0.87
	SSQ	0.1020	0.0880	0.0855
	T	0.84	0.01	0.45
	U	0.88	1.25	0.77
8	V	1.07	0.26	0.39
	W	0.00	0.52	0.68
	SSQ	0.1160	0.1458	0.1099
	T	0.93	0.68	0.00
	U	0.86	0.82	1.20
16	V	1.20	1.04	0.27
	W	0.00	0.00	0.56
	SSQ	0.1272	0.1463	0.1586

CONCLUDING REMARKS

This paper has raised a number of key modelling issues, but it has not pursued any of them to a conclusion. We have just seen that, whatever the view one may take of the deterministic/continuous versus stochastic/discrete issue at the mini-battle level, there must at least be some doubt that a Lanchester formulation will serve as a good summary descriptor of mini-battle dynamics. The priority now must surely be to produce an adequate account of mini-battle formation. This in turn will provide one necessary precondition for addressing the all-important issues raised under the subheading of target acquisition. When this has been done we can commence the task of laboratory experimentation in order to gain a better understanding of the dynamics of the mobile land engagement. We can then make some attempt to allow for the effects of combat degradation, and can hope to define a satisfactory logic for deciding at what point an engagement may be considered at an end.

Accordingly, the next paper in this series will examine in greater depth the topics of mini-battle formation and target acquisition. The following paper will go on to consider combat degradation and modelling criteria for victory or defeat. In each case the approach will be the same. For each of the four subheadings there will be a review of the evidence from trials and war records, and then a modelling approach will be proposed which is consistent with this evidence. Model components built to these specifications will then be assembled into a description of the dynamics of the mini-battle and of the engagement as a whole. Finally, laboratory experiments will, among other things, link the outputs of this schematic model to possible Lanchester formulations.

REFERENCES

Ancker, C.J.Jr. and Gafarian, A.V. 1988. The Validity of Assumptions Underlying Current Uses of Lanchester Attrition Rates. US Army TRADOC Analysis Command, White Sands Missile Range, NM, TRAC-WSMR-TD-7-88.

Brackney, H. 1959. The Dynamics of Military Combat. *Opns.Res.*, 7, 30-44.

Brown, R.H. 1963. Theory of Combat: The Probability of Winning. *Opns.Res.*, 11, 418-425.

Cooke, K.J., Stanley, P.A. and Hinton, J.L. 1994. The ORACLE Approach to Target Acquisition and Search Modelling. In Peli, E. (Ed.) *Applied Spatial Vision Models for Target Detection.* World Scientific Publishing Inc., River Edge, NJ

Goldie, C.M. 1977. Lanchester Square-Law Battles: Transient and Terminal Distributions. *J.Appl.Prob.*, 14, 604-610.

Hartley, D.S. 1991. Confirming the Lanchestrian Linear-Logarithmic Model of Attrition. Martin Marietta Energy Systems, Inc, Oak Ridge, TN, *K/DSRD-263/RI.*

Helmbold, R.L. 1961a. Lanchester Parameters for some Battles of the Last Two Hundred Years. Combat Operations Research Group, Fort Belvoir, Va., CORG-SP-122.

Helmbold, R.L. 1961b. Historical Data and Lanchester's Theory of Combat. Combat Operations Research Group, Fort Belvoir, Va., CORG-SP-128.

Helmbold, R.L. 1964. Historical Data and Lanchester's Theory of Combat, Part II. Combat Operations Research Group, Fort Belvoir, Va., CORG-SP-190.

Helmbold, R.L. 1965. A Modification of Lanchester's Equations. *Opns.Res.*, 13, 857-859.

Helmbold, R.L. 1993. Combat Analysis. Article series, *Phalanx*, 26, 23 et seq.

Hinton, J.L. 1992. The BAe SRC Visual Performance Model ORACLE: An Overview - 1992. Sowerby Research Centre, British Aerospace, Filton, Bristol, *Report No. JS 12138.*

Lanchester, F.W. 1916. *Aircraft in Warfare: The Dawn of the Fourth Arm.* Constable, London.

McNaught, K.R. 1995. The Effects of Nodal Structure on Exponential Stochastic Lanchester Battles. Royal Military College of Science, Shrivenham. Paper awaiting publication.

Marshall, S.L.A. 1947. *Man Against Fire: The Problem of Battle Command in Future War.* Morrow & Co, New York.

Osipov, M. 1915. The Influence of the Numerical Strength of Engaged Forces on their Casualties. (Originally published in the Tzarist Russian Journal, Military Collection, June-October.) *transl.* Helmbold, R.L. and Rehm, A.S. 1991. U.S.Army Concepts Analysis Agency, *Research Paper* CAA-RP-91-2.

Overington, I. 1982. Towards a Complete Model of Photopic Visual Threshold Performance. *Opt.Engng.*, 21, 002-013.

Rowland, D. 1984. Field Trials and Modelling. Paper Delivered to International Symposium on Advances in Combat Modelling, Royal Military College of Science, Shrivenham.

Smith, D.G. 1965. The Probability Distribution of the Number of Survivors in a Two-Sided Combat Situation. *Opnl.Res.Quart.*, 16, 429-437.

Speight, L.R. 1995. Back to Basics: Lanchester's Equations and the Mini-Engagement in the Mobile Land Battle. Royal Military College of Science, Shrivenham. *AMOR 95/3.*

Taylor, J.G. 1980. *Force-on-Force Attrition Modelling.* Military Applications Section, Operations Research Society of America.

Taylor, J.G. 1983. *Lanchester Models of Warfare, Vols I & 2.* Military Applications Section, Operations Research Society of America.

Watson, R.K. 1967. An Application of Martingale Methods to Conflict Models. *Opns.Res.*, 24, 380-382.

Weale, T.G. 1992. Two Numerical Methods for Computing the Probability of Outcome of a Battle of Lanchester Type. *J.Opl.Res.Soc.*, 43, 797-807.

Weiss, G.H. 1963. Comparison of a Deterministic and a Stochastic Model for Interaction Between Antagonistic Species. *Biometrics*, 19, 595-602.

Weiss, H.K. 1966. Combat Models and Historical Data: The US Civil War. *Opns. Res.*, 14, 759-789.

Willard, D. 1962. *Lanchester as a Force in History: An Analysis of Land Battles of the Years 1618-1905.* McLean, Va: Research Analysis Corp.

Editorial Policy and Submission of Papers

MILITARY OPERATIONS RESEARCH is published by the Military Operations Research Society, in association with the Military Applications Section of the Institute for Operations Research Society of America and Management Science.

The aim of the journal is to develop, promote, and the coordinate the science and practice of military operations research; to establish channels of communications that link government, industry, and academia; and to facilitate the interchange of ideas among practitioners, academics, and policy makers.

EDITORIAL POLICY:

Military Operations Research will publish invited and submitted original papers, review papers, case studies, and notes and comments on all aspects of military operations research. Papers of an expository or tutorial nature that are of wide interest will be published. However, *MOR* will not publish papers that are accessible or interesting only to a small group of specialists. Papers on the history of OR are welcome provided they convey the insights of current interest. Of particular interest are papers that present case studies showing innovative applications of OR; papers that introduce interesting new problem areas; papers on educational issues; and papers that relate policy issues to operations research. Commentaries on papers published in *MOR* will be encouraged; authors will have an opportunity to respond to comments before the entire dialogue is published. Papers suitable for *MOR* should be readable by those who are comfortable with a level of mathematics appropriate to a masters program in operations research. All articles will be reviewed. *MOR* reserves the right to refuse any submission, invited or not, and to make any editorial modifications needed before publication.

SUBMISSION OF PAPERS:

Manuscripts (3 copies) should be submitted directly to the Editor, as follows:

Editor, *MOR*
Military Operations Research Society
101 South Whiting Street, Suite 202
Alexandria, VA 22304

All submissions must be unclassified and be accompanied by release statements where appropriate. By submitting a paper for review, an author certifies that the manuscript has been cleared for publication, is not copyrighted, has not been accepted for publication in any other publication, and is not under review elsewhere at the same time. All authors will be required to sign a copyright agreement with MORS (see following page). For paper layout and style, authors should refer to a recent issue of the journal.

MILITARY OPERATIONS RESEARCH SOCIETY, INC.

COPYRIGHT ASSIGNMENT AGREEMENT

THIS AGREEMENT, made by and between ____________________________, hereinafter ("Assignor(s)") and MILITARY OPERATIONS RESEARCH SOCIETY, INC., hereinafter Assignee).

Assignor(s) (*for U.S. Government employees: to the extent legally permissible*) hereby grants, transfers, assigns, and conveys to Assignee, its successors and assigns, the entire title, right, interest, ownership and all subsidiary rights in and to an article entitled ____

____

____(hereinafter the "Work"), including, but not limited to, the right to secure copyright registration of the Work and to any resulting registration in Assignee's name as claimant. Notwithstanding the foregoing, Assignor(s) reserves all proprietary rights other than copyright, such as patent rights; a royalty-free right to reproduce the Work; and the right to prepare derivative works based upon the copyright in the Work.

Assignor(s) hereby confirms that Assignor(s) owns the entire title, right and interest in the work, including the right to reproduce, prepare derivative works based upon the copyright in the Work, and distribute the Work, whether or not the Work constitutes a "work made for hire" as defined in 17 U.S.C. Section 201(b); Assignor(s) agrees that no rights in the Work are retained by Assignor(s) except as stated above. Assignor(s) agrees to execute any documents that might be necessary to perfect Assignee's ownership of copyrights in the Work and to registration.

Assignor(s) represents that the Work has not been copyrighted or published; that it is not being submitted for publication elsewhere; and, if the Work is officially sponsored by the U.S. Government, that it has been approved for open publication.

This Agreement constitutes the entire agreement between the parties hereto; this Agreement supersedes any prior oral or written agreement or understanding between the parties; and, **in the case of a Work Made for Hire, this Agreement has been signed by the Assignor(s)'s employer.** This Agreement shall only be effective if and upon the date that the Work is accepted by Assignee for publication in *Military Operations Research*, the Journal of the Military Operations Research Society or *PHALANX*, the Bulletin of the Military Operations Research Society.

ASSIGNOR(S)

By:__________________________

Title:________________________

Date:________________________

ASSIGNOR(S)

By:__________________________

Title:________________________

Date:________________________

MILITARY OPERATIONS RESEARCH SOCIETY, INC. (ASSIGNEE)

By:__________________________

Title:________________________

Date:________________________